

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2001, 15:09:55 ; Search time 19.72 Seconds

(without alignments)
626.129 Million cell updates/sec

Title: US-09-602-597-4

Perfect score: 1101
Sequence: 1 MAEPSPKSTRTSSPAQGA.....GSOSRVTLQASGALRGVSP 216

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq_0401:*

1: /SID56/gcgdata/geneseq/geneseqp/AA1980.DAT:*

2: /SID56/gcgdata/geneseq/geneseqp/AA1981.DAT:*

3: /SID56/gcgdata/geneseq/geneseqp/AA1982.DAT:*

4: /SID56/gcgdata/geneseq/geneseqp/AA1983.DAT:*

5: /SID56/gcgdata/geneseq/geneseqp/AA1984.DAT:*

6: /SID56/gcgdata/geneseq/geneseqp/AA1985.DAT:*

7: /SID56/gcgdata/geneseq/geneseqp/AA1986.DAT:*

8: /SID56/gcgdata/geneseq/geneseqp/AA1987.DAT:*

9: /SID56/gcgdata/geneseq/geneseqp/AA1988.DAT:*

10: /SID56/gcgdata/geneseq/geneseqp/AA1989.DAT:*

11: /SID56/gcgdata/geneseq/geneseqp/AA1990.DAT:*

12: /SID56/gcgdata/geneseq/geneseqp/AA1991.DAT:*

13: /SID56/gcgdata/geneseq/geneseqp/AA1992.DAT:*

14: /SID56/gcgdata/geneseq/geneseqp/AA1993.DAT:*

15: /SID56/gcgdata/geneseq/geneseqp/AA1994.DAT:*

16: /SID56/gcgdata/geneseq/geneseqp/AA1995.DAT:*

17: /SID56/gcgdata/geneseq/geneseqp/AA1996.DAT:*

18: /SID56/gcgdata/geneseq/geneseqp/AA1997.DAT:*

19: /SID56/gcgdata/geneseq/geneseqp/AA1998.DAT:*

20: /SID56/gcgdata/geneseq/geneseqp/AA1999.DAT:*

21: /SID56/gcgdata/geneseq/geneseqp/AA2000.DAT:*

22: /SID56/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	930	84.5	182	Y70458	Human membrane cha
2	848	77.0	182	R99799	NTII-11 nerve prot
3	580	52.7	113	B34711	Human secreted pro
4	325	29.5	173	V53643	A bone marrow secr
5	307	27.9	153	W52117	Human proteolipid
6	191	17.3	101	Y41391	Human secreted pro
7	148.5	13.5	169	B56719	Human prostatic can
8	147.5	13.4	265	B49502	Clone HYASC03. HO
9	144	13.1	153	P81879	Sequence of full-1
10	144	13.1	153	P80929	Sequence of human
11	137.5	12.5	276	B42947	Human ORFX ORF2711

12	137.5	12.5	322	21	B15549	Human immune syste
13	137.5	12.5	322	21	B19729	Human SECC Clone 4
14	137.5	12.5	322	21	Y99349	Human PRO110 (UNG)
15	137	12.4	254	21	B56922	Human prostate can
16	127.5	11.6	254	21	B24446	Human secreted pro
17	127.5	11.6	299	21	B23042	Human myeloid upre
18	115	10.4	175	21	B29644	Human membrane-ans
19	110.5	10.0	156	20	Y33297	Human membrane spa
20	94	8.5	34	20	Y41565	Fragment of human
21	92.5	8.4	125	21	B42930	Human ORFX ORF2694
22	91	8.3	421	20	Y35514	Beta-adrenergic re
23	86.5	7.9	400	14	R32501	Human ORFX ORF2170
24	85	7.7	1088	21	B42406	Human ORFX ORF2170
25	84.5	7.7	246	21	B29615	Cat flea HMT NMDA
26	82	7.4	456	19	W69555	Human equilibrativ
27	82	7.4	456	21	B15520	Human ENTI protein
28	81.5	7.4	388	13	R25698	Murine adrenergic
29	81.5	7.4	400	15	R54992	Murine beta-3 adre
30	81	7.4	230	21	G19552	Arabidopsis thalia
31	81	7.4	237	21	G19551	Arabidopsis thalia
32	81	7.4	255	21	B19414	An aromatic class
33	81	7.4	272	21	G19550	Arabidopsis thalia
34	81	7.4	553	20	Y38796	Neisseria meningit
35	80.5	7.3	359	21	Y69485	Amino acid sequenc
36	80	7.3	232	21	G32159	Arabidopsis thalia
37	80	7.3	239	21	G32158	Arabidopsis thalia
38	80	7.3	280	21	G32157	Arabidopsis thalia
39	79.5	7.2	405	19	W44933	Canine beta-3 adre
40	79.5	7.2	476	18	W24089	Balanus amphitrite
41	79	7.2	306	21	B43004	Human ORFX ORF2768
42	79	7.2	306	22	B49652	Human SEC4 protein
43	79	7.2	394	21	Y81725	Streptococcus pneu
44	79	7.2	1670	21	G31673	Arabidopsis thalia
45	79	7.2	1753	21	G31672	Arabidopsis thalia

ALIGNMENTS

RESULT	1
ID	Y70458 standard; Protein; 182 AA.
XX	Y70458:
AC	
XX	
DT	21-JUN-2000 (first entry)
XX	
DE	Human membrane channel protein-8 (MECHP-8).
XX	
KW	Membrane channel protein-8; MECHP-8; diagnosis; treatment; lymphoma;
KW	cell proliferative disorder; bursitis; atherosclerosis; cancer; sarcoma;
KW	inflammatory disorder; AIDS; Addison's disease; cystic fibrosis; asthma;
KW	diabetes mellitus; osmoregulatory disorder; diarrhoea; renal failure;
KW	muscular disorder; myocarditis; Duchenne's muscular dystrophy; nocturnal;
KW	cardiovascular disorder; hypertension; bronchitis; vasculitis; cardiac;
KW	neurological disorder; Alzheimer's disease; Parkinson's disease; human;
KW	Huntington's disease; antiarteriosclerotic; hepatotropic; cytotoxic;
KW	anti-HIV; antianaemic; neuroprotective; immunomodulator; antidiabetic;
KW	hypotensive; vasotrophic; antisthmatic; antiinflammatory; antidepressant;
KW	anticonvulsant; thrombolytic; antiParkinsonian; immunostimulant.
XX	
OS	Homo sapiens.
XX	
Key	Location/Qualifiers
FT	Modified-site 9
FT	/note- "Phosphorylation site"
FT	Modified-site 130
FT	/note- "Phosphorylation site"
FT	Modified-site 108
FT	/note- "Glycosylation site"
FT	/note- "Glycosylation site"
FT	Domain 36..55
FT	/label- Transmembrane_domain
FT	Domain 65..84

```

FT      Domain
FT      /label= Transmembrane_domain
FT      103..130
FT      /label= Transmembrane_domain
XX
XX      WO200012711-A2.
XX
XX      09-MAR-2000.
XX
XX      02-SEP-1999; 99WO-US20468.
XX
XX      02-SEP-1998; 98US-0145815.
XX      PR      12-NOV-1998; 98US-0191283.
XX      PR      09-DEC-1998; 98US-0208821.
XX      PR      26-JAN-1999; 99US-0237506.
XX      PR      10-FEB-1999; 99US-0247891.
XX
XX      (INCY-) INCYTE PHARM INC.
XX
XX      Au-Young J, Bandman O, Tang YT, Reddy R, Hillman JL, Yue H;
XX      Lal P, Corley NC, Guegler KJ, Gorgone G, Baughn MR, Azimzal Y;
XX      WPI; 2000-256643/22.
XX      N-PSDB; 251624.
XX
XX      Novel human membrane channel protein and polynucleotide useful for
XX      diagnosing and creating cell proliferative, inflammatory, secretory,
XX      osmoregulatory, muscular, cardiovascular and neurological disorders -
XX
XX      Claim 1; Page 107; 140pp; English.
XX
XX      The present sequence is the human membrane channel protein-8 (MECHP-8),
XX      which is expressed in nervous and gastrointestinal tissues. Anti-MECHP
XX      antibodies can be used as therapeutic antagonists and reagents for
XX      diagnosis and monitoring diseases. MECHP cDNA can be used for diagnosis
XX      of MECHP-related diseases and gene mapping. MECHP can be used for
XX      treatment of cell proliferative disorders such as bursitis and
XX      atherosclerosis, cancers like lymphoma and sarcoma, inflammatory
XX      disorders like AIDS and Addison's disease, transport/secretory disorders
XX      like cystic fibrosis and diabetes mellitus, osmoregulatory disorders like
XX      diarrhoea and renal failure, muscular disorders like myocardiitis and
XX      Duchenne's muscular dystrophy, cardiovascular disorders like hypertension
XX      and vasculitis, congenital lung anomalies like bronchitis and asthma and
XX      neurological disorders like Alzheimer's disease, Parkinson's disease and
XX      Huntington's disease.
XX
XX      Sequence 182 AA;
XX
XX      Query Match 84.5%; Score 930; DB 21; Length 182;
XX      Best Local Similarity 100.0%; Pred. No. 6; Le-98;
XX      Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX
XX      1 MAEPSPKYSRTSPPOGAASVSAIRPDLGVRSHGLMIMLQVGLWALIAATPY 60
XX      Db      1 maeipksrctstspagaasvsairpdlgtvrsrlgalmllqvgllwailatpy 60
XX
XX      61 HLYPAYGWFVAVFVLMVTVLVFNLYLQILMKIYMPMPVLYIMFISATVYITAFI 120
XX      Db      61 hlypaygwmfvavfvlmvtvlyvfnlylqlmkllympmpvlylmfistvtlyitafi 120
XX
XX      121 AGSAANDVLSLGRTPRYNQRAAASPACLVMTAIVSAFSTYQAMRGVGSNAATSQMAG 180
XX      Db      121 acsaavdlslgrtprynpqraaasfacclvmiaiyvsafstygawrgvgsnaatsqmag 180
XX
XX      181 YA 182
XX      181 ya 182
XX
XX      RESULT 2
XX      ID R99799 standard: Protein: 182 AA.
XX

```

```

AC. R99799; 02-JAN-1997 (first entry)
XX
XX
DE NTII-11 nerve protein, facilitates regeneration of nerve cells.
XX
XX
KM Probe; central nervous system; peripheral nervous system; CNS; PNS;
XX axon; neurone; neuronal cells; glial cells; Schwann cells; trauma;
XX pathology; regeneration; MS2; metalloproteinase.
XX
XX OS Homo sapiens.
XX
XX PN MO6617865-A2.
XX
XX PD 13-JUN-1996.
XX
XX PF 05-DEC-1995; 95WO-EP04777.
XX
XX PR 27-JAN-1995; 95DE-4002525.
XX PR 05-DEC-1994; 94DE-4443159.
XX
XX PA (BOE) BOEHRINGER MANNHEIM GMBH.
XX
XX PI Gillen C, Gleichmann M, Mueller H;
XX
XX DR WPI: 1996-287114/29.
XX DR N-PSDB; T34592.
XX
XX PT Nucleic acids associated with damaged or regenerating nerve cells
XX PT for diagnostic or therapeutic use.
XX
XX PS Claim 4; Page 45-46; 54pp; German.
XX
XX CC Nucleic acids described in T34587-T34616 are useful as probes for
XX CC detecting traumatic and pathological changes in the peripheral
XX CC nervous system. They may also be used to identify hybridising
XX CC sequences from an ischial nerve cDNA library, these sequences encode
XX CC proteins expressed in damaged and/or regenerating nerve cells. They
XX CC can be used to transfect cells which then express the protein which
XX CC can then be harvested and studied. The nucleic acids are preferably
XX CC selected from NTII-11 which corresponds to plasmolipin cDNA with a
XX CC 234 bp N-terminal extension and is expressed in glial cells of the
XX CC central nervous system and Schwann cells of the peripheral nervous
XX CC system. NTII-1, which codes for this protein having 79% homology
XX CC with human C9orf41 in three domains and may be involved in regulating
XX CC Schwann cell growth as well as being therapeutically useful for
XX CC regenerating nerve cells; and CRTI-7 which codes for a protein
XX CC which is similar to the macrophage-specific cysteine-rich mouse
XX CC protein MS2 and a metalloproteinase and which inhibits interaction
XX CC between Schwann cells and the extracellular matrix and is useful
XX CC for the therapy of nerve damage by facilitating regeneration of
XX CC damaged axons.
XX
XX Sequence 182 AA:
XX
XX
XX Query Match 77.0%; Score 848; DB 17; Length 182;
XX Best Local Similarity 88.5%; Pred. No. 1,3e-88;
XX Matches 161; Conservative 13; Mismatches 8; Indels 0; Gaps 0;
XX
XX 1 MAEFSKYSTRTSSPAQGAASVSLRPDLGFRSLGALMLDLQVLGLLWALLADPPY 60
XX 1 maefskvstsrtspsagvgaaesamrpdlgfrsalgylallqvlvgllwalladp 60
XX
XX 61 HLPRAYGVWMPAYAVLMLVTVLLEMLYLFOFLMKIKYMPMPPLVLMFNISAVLYITAFI 120
XX 61 hlpraygvwmfayavllvltvlvllfllfqlhmkikymppvllvlyfvaavlyitafv 120
XX
XX 121 ACSAAVDTLSLGRTPYNOBAAAFACIWMATYGVSAFFSYQANRGVGSNNAATSGMAG 180
XX 121 acaaaavdtlsigrtpynqrsaasifacimlatyglalsalfsqawrgvgsnaatsqmag 180
XX
XX 181 YA 182

```



```
XX 23-SEP-1999.
PD
XX
XX 18-MAR-1999; 99W0-US05804.
XX
XX 19-MAR-1998; 98US-0078563.
XX 19-MAR-1998; 98US-0078566.
XX 19-MAR-1998; 98US-0078573.
XX 19-MAR-1998; 98US-0078574.
XX 19-MAR-1998; 98US-0078576.
XX 19-MAR-1998; 98US-0078577.
XX 19-MAR-1998; 98US-0078578.
XX 19-MAR-1998; 98US-0078579.
XX 19-MAR-1998; 98US-0078581.
XX 01-APR-1998; 98US-0080312.
XX 01-APR-1998; 98US-0080313.
XX 01-APR-1998; 98US-0080314.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Ni J, Rosen CA, Yu G, Young PE, Feng P, Soppet DR;
PI Wei Y, Endress GA, Duan RD, Kyaw H, Ebner R, Lafleur DW;
PI Olsen HS, Shi Y, Moore PA;
XX
XX MPI: 1999-562050/47.
XX N-PSDB; Z24894.
XX
XX New isolated human genes, useful for diagnosis and treatment of e.g.
XX cancers, neurological disorders, immune diseases, inflammation or blood
XX disorders -
XX
XX Claim 11; Page 407408; 484pp; English.
XX
XX This sequence represents a secreted human protein encoded by the gene
XX clone detailed in the descriptor line. The gene can be used to generate
XX fusion proteins by linking to the gene to a human immunoglobulin Fc
XX portion (e.g. Z24802) for increasing the stability of the fused protein
XX as compared to the human protein only.
XX The invention relates to 95 novel genes and their fragments (nucleic
XX acid sequences: Z24811-Z24907; amino acid sequences Y41308-Y41404) which
XX are useful for preventing, treating or ameliorating medical conditions
XX e.g. by protein or gene therapy. Also, pathological conditions can be
XX diagnosed by determining the amount of the new polypeptides in a sample
XX or by determining the presence of mutations in the new polynucleotides.
XX Specific uses are described for each of the 95 polynucleotides, based on
XX which tissues they are most highly expressed in (see Z24811 for described
XX uses).
XX
XX Sequence 101 AA;
XX
XX Query Match 17.3%; Score 191; DB 20; Length 101;
XX Best Local Similarity 41.0%; Pred. No. 3.4e-14;
XX Matches 41; Conservative 20; Mismatches 31; Indels 8; Gaps 2;
XX
XX 70 MRYAVFLMTVLTLENYLFQLMKLYWPPWPLVIMFINSATVLYTAFIACSAVDLT 129
XX ||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
XX 1 mfvavfvlvtvfllytmtcytrtpyptwtvgjcfngsafvlylsaa-----avvdaa 55
XX
XX 130 SL---RGTRPYNORAAAFACIYMYAGVSAAFFSQAMR 166
XX || : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
XX 56 svspkcdshnfmsaasffafvlycgnlyfsfxawr 95
XX
XX RESULT 7
XX ID B56719 standard; Protein; 169 AA.
XX AC B56719;
XX XX
XX DT 13-MAR-2001 (first entry)
XX DE Human prostate cancer antigen protein sequence SEQ ID NO:1297.
```

```
XX
XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
XX neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
XX vulnary; gastrointestinal; nephrotoxic; antiinfective; gynaecological;
XX antibacterial; gene therapy; neural; immune; reproductive; renal;
XX gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX wound; infectious disease.
XX
XX Homo sapiens.
XX
XX WO20005174-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000W0-US05988.
XX
XX 12-MAR-1999; 99US-0124270.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
XX
XX Rosen CA, Ruben SM;
XX
XX MPI: 2000-587513/55.
XX N-PSDB; F15922.
XX
XX prostate cancer associated gene sequences, referred to as prostate
XX cancer antigens, useful for treatment, prevention, and diagnosis of
XX disorders such as prostate cancer -
XX
XX Claim 11; Page 1718; 2338pp; English.
XX
XX F15566 to F16505 encode the human prostate cancer associated proteins,
XX called prostate cancer antigens, given in B56363 to B57302. The prostate
XX cancer antigens can have neuroprotective, cytostatic, cardioactive,
XX immunomodulatory, muscular, vulnary, gastrointestinal, nephrotoxic,
XX antiinfective, gynaecological and antibacterial activities, and can be
XX used in gene therapy. The prostate cancer antigen polynucleotides may be
XX used for detection of prostate cancer. Chromosome identification, as
XX chromosome markers, and for numerous other diagnostic or research
XX purposes. The prostate cancer antigens may be used to treat disorders
XX such as neural, immune, muscular, reproductive, gastrointestinal,
XX pulmonary, cardiovascular, renal, and proliferative disorders, wounds,
XX and infectious diseases. F16506 to F16514 to B57303 represent sequences
XX used in the exemplification of the present invention.
XX
XX Sequence 169 AA;
XX
XX Query Match 13.5%; Score 148.5; DB 21; Length 169;
XX Best Local Similarity 30.2%; Pred. No. 4.7e-09;
XX Matches 38; Conservative 31; Mismatches 52; Indels 5; Gaps 4;
XX
XX 39 AMLTQIVLGLVYMALIADTPHYLYP-AYGWWVFVFLVYVFNLYTLQHLKLM 97
XX || : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
XX 45 atfipellfgflwtnvaat-hlvypdlqgwwyvafstlsimllsylygfykr--f 101.
XX
XX 98 VPMPLVIMFINSATVLYTR-AFIACSAVDLTSLRGTRPYNORAAAFACIYMYAGV 156
XX || : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
XX 102 esvrvldslsyhgttglylmsaavlyghatvseklldiprlylmsaasffafvlyl 161
XX
XX 157 SAFFSY 162
XX |||
XX 162 hafsly 167
XX
XX RESULT 8
XX ID B49502 standard; Protein; 265 AA.
XX AC B49502;
XX XX
XX DT 09-MAR-2001 (first entry)
```

XX	Cione HYAC03.	
DE		
XX		
KW	Human; 4 transmembrane superfamily receptor protein;	
KW	endocrine; cardiovascular; cerebrovascular disease; neural disorder;	
KW	reproductive; skin; renal system; autoimmune; hyperproliferative; ocular;	
KW	bacterial infection; viral; fungal.	
XX		
OS	Homo sapiens.	
XX		
PN	W0200070076-A1.	
XX		
PD	23-NOV-2000.	
XX		
PF	18-MAY-2000; 2000WO-US13504.	
XX		
PR	19-MAY-1999; 99US-0135122.	
PR	03-JUN-1999; 99US-0137797.	
PR	11-JUN-1999; 99US-0138573.	
PR	18-AUG-1999; 99US-0149447.	
PR	28-JAN-2000; 2000US-0178770.	
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Ruben SM, Ni J, Fan P, Roschke V, Shi Y, Komatsoulis GA;	
PI	Rosen CA;	
XX		
DR	WPI: 2001-007502/01.	
DR	N-PSDB; C90012.	
XX		
PT	Isolated nucleic acid molecule encoding human soluble 4 transmembrane	
PT	superfamily receptor protein, useful for diagnosing, treating and/or	
PT	preventing disorders e.g. Alzheimer's, cancer and arrhythmia.	
XX		
PS	Claim 11; Pages 278-279; 297pp; English.	
XX		
CC	The present invention relates to isolated nucleic acids and proteins	
CC	encoding human soluble 4 transmembrane superfamily receptor protein (see	
CC	C90012-C90023 and B49502-B49513). The present sequence is one such	
CC	protein. The present protein can be used to screen for binding partners	
CC	and molecules which modify its activity. Antibodies specific for the	
CC	present protein can be used to treat and/or prevent diseases associated	
CC	with aberrant expression or activity of the present protein e.g.	
CC	endocrine disorders e.g. Addison's disease, (cardio)vascular diseases	
CC	e.g. arrhythmia and atherosclerosis, cerebrovascular diseases, neural	
CC	disorders e.g. Alzheimer's and Parkinson's disease, reproductive	
CC	disorders, skin disorders e.g. psoriasis, renal system disorders e.g.	
CC	nephritis, (auto)immune system disorders e.g. graft vs. host disease,	
CC	hyperproliferative disorders e.g. neoplasms of the pancreas, ocular	
CC	disorders e.g. glaucoma and infections caused by bacteria, viruses and	
XX	fungi.	
XX		
SQ	Sequence 265 AA;	
	Query Match 13.4%; Score 147.5; DB 22; Length 265;	
	Best Local Similarity 29.2%; Pred. No. 1.1e-08;	
	Matches 50; Conservative 32; Mismatches 56; Indels 23; Gaps 7	
QY	2 AEPFSKVTSTSPQAQAEASVSALRPDLGFRGRGLGALMLQLVLLYVLIADTPYH 61	
Db	5 assyppttepsqgrq---lagircdpdyrgalgrkvaqvilalaficici-etima 59	
QY	62 LYPAYG--WYMFVAVFLMWTIVLENLVLFQIHMKLYMVPWPLVMTFNISATVLYTAF 119	
Db	60 cpsceglyfeyfscsfayvtgvtgllmfslnlhmrpqinwnltdly-ntg----lsaf 113	
QY	120 IACSAVDLTSLRGTRPNORA-----AASFFACLVMITAYGVSAFFSQANR 166	
Db	114 iffiasivlaal-----nhrgaiaaavifgfataayavntflavqkvr 158	

RESULT 9

CC downregulated in astrocytoma. Therefore restoring its expression
CC could be useful for the treatment of astrocytomas. Since it is
CC a multiple membrane spanning protein, gene therapy would be
CC necessary.

SO Sequence 322 AA;

Query Match 12.5%; Score 137.5; DB 21; Length 322;
Best Local Similarity 25.6%; Pred. No. 2e-07;
Matches 45; Conservative 29; Mismatches 79; Indels 23; Gaps 4;

QY 9 STPTSSPAQGAESVALRPDLGFVNSRLGALMLQVLGLVALLADPPHYLPAYGW 68
D 12 tttsssglsgpmlygspra---lcpqlglrlrlqlvscvafslvasvgawtsgmgnw 67
QY 69 VMFVAFLMVTIVLNLV-----FOLMKILMVPMLVIMFNLSAVLYTATAC 122
D 68 smtfcvctsvlllllvefcglqarfpiswrifpftfacyaaftclsasliypetyv-- 125
QY 123 SAAVDLTSLNGTRPYNORAAASFACILMIAYGVSAFFSYOAKGVGNATASOMA 178
D 126 -----qflshgtrchalaatftfsciacvayatev-----awtrarpgetlgyma 170

RESULT 14

ID Y99349 standard; Protein; 322 AA.

AC Y99349;

DT 08-AUG-2000 (first entry)

DE Human PRO1110 (UNQ53) amino acid sequence SEQ ID NO:31.

KW Human: PRO polypeptide; membrane bound protein; receptor; diagnosis;
transmembrane; secretion; immunoadhesion; pharmaceutical; screening.

OS Homo sapiens.

PN WO200012708-A2.

PD 09-MAR-2000.

PF 01-SEP-1999; 99WO-US20111.

XX 01-SEP-1998; 98US-0098716.
PR 01-SEP-1998; 98US-0098749.
PR 01-SEP-1998; 98US-0098750.
PR 02-SEP-1998; 98US-0098803.
PR 02-SEP-1998; 98US-0098821.
PR 02-SEP-1998; 98US-0098843.
PR 09-SEP-1998; 98US-0099336.
PR 09-SEP-1998; 98US-0099396.
PR 09-SEP-1998; 98US-0099598.
PR 09-SEP-1998; 98US-0099602.
PR 09-SEP-1998; 98US-0099642.
PR 10-SEP-1998; 98US-0099741.
PR 10-SEP-1998; 98US-0099754.
PR 10-SEP-1998; 98US-0099763.
PR 10-SEP-1998; 98US-0099792.
PR 10-SEP-1998; 98US-0099808.
PR 10-SEP-1998; 98US-0099812.
PR 10-SEP-1998; 98US-0099815.
PR 15-SEP-1998; 98US-0100385.
PR 15-SEP-1998; 98US-0100388.
PR 15-SEP-1998; 98US-0100390.
PR 16-SEP-1998; 98US-0100584.
PR 16-SEP-1998; 98US-0100627.
PR 16-SEP-1998; 98US-0100661.
PR 16-SEP-1998; 98US-0100662.
PR 16-SEP-1998; 98US-0100664.

PR 17-SEP-1998; 98US-0100683.
PR 17-SEP-1998; 98US-0100684.
PR 17-SEP-1998; 98US-0100710.
PR 17-SEP-1998; 98US-0100711.
PR 17-SEP-1998; 98US-0100919.
PR 17-SEP-1998; 98US-0100930.
PR 18-SEP-1998; 98US-0100848.
PR 18-SEP-1998; 98US-0100849.
PR 18-SEP-1998; 98US-0101014.
PR 18-SEP-1998; 98US-0101068.
PR 18-SEP-1998; 98US-0101071.
PR 22-SEP-1998; 98US-0101279.
PR 23-SEP-1998; 98US-0101471.
PR 23-SEP-1998; 98US-0101472.
PR 23-SEP-1998; 98US-0101474.
PR 23-SEP-1998; 98US-0101475.
PR 23-SEP-1998; 98US-0101476.
PR 23-SEP-1998; 98US-0101477.
PR 23-SEP-1998; 98US-0101479.
PR 24-SEP-1998; 98US-0101738.
PR 24-SEP-1998; 98US-0101741.
PR 24-SEP-1998; 98US-0101743.
PR 24-SEP-1998; 98US-0101915.
PR 24-SEP-1998; 98US-0101916.
PR 29-SEP-1998; 98US-0102207.
PR 29-SEP-1998; 98US-0102240.
PR 29-SEP-1998; 98US-0102307.
PR 29-SEP-1998; 98US-0102330.
PR 29-SEP-1998; 98US-0102331.
PR 30-SEP-1998; 98US-0102484.
PR 30-SEP-1998; 98US-0102487.
PR 30-SEP-1998; 98US-0102570.
PR 30-SEP-1998; 98US-0102571.
PR 01-OCT-1998; 98US-0102684.
PR 01-OCT-1998; 98US-0102687.
PR 02-OCT-1998; 98US-0102965.
PR 06-OCT-1998; 98US-0103258.
PR 06-OCT-1998; 98US-0103449.
PR 07-OCT-1998; 98US-0103314.
PR 07-OCT-1998; 98US-0103315.
PR 07-OCT-1998; 98US-0103328.
PR 07-OCT-1998; 98US-0103395.
PR 07-OCT-1998; 98US-0103396.
PR 07-OCT-1998; 98US-0103401.
PR 08-OCT-1998; 98US-0103633.
PR 08-OCT-1998; 98US-0103678.
PR 08-OCT-1998; 98US-0103679.
PR 08-OCT-1998; 98US-0103711.
PR 14-OCT-1998; 98US-0104257.
PR 20-OCT-1998; 98US-0104987.
PR 20-OCT-1998; 98US-0105000.
PR 20-OCT-1998; 98US-0105002.
PR 21-OCT-1998; 98US-0105104.
PR 22-OCT-1998; 98US-0105169.
PR 22-OCT-1998; 98US-0105266.
PR 26-OCT-1998; 98US-0105693.
PR 26-OCT-1998; 98US-0105807.
PR 27-OCT-1998; 98US-0105881.
PR 27-OCT-1998; 98US-0105882.
PR 27-OCT-1998; 98US-0106062.
PR 28-OCT-1998; 98US-0106023.
PR 28-OCT-1998; 98US-0106029.
PR 28-OCT-1998; 98US-0106030.
PR 28-OCT-1998; 98US-0106032.
PR 28-OCT-1998; 98US-0106033.
PR 28-OCT-1998; 98US-0106178.
PR 29-OCT-1998; 98US-0106248.
PR 29-OCT-1998; 98US-0106384.
PR 29-OCT-1998; 98US-0106500.
PR 30-OCT-1998; 98US-0106464.
PR 03-NOV-1998; 98US-0106856.
PR 03-NOV-1998; 98US-0106902.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2001, 15:09:55 ; Search time 14.22 Seconds
(without alignments)
1043.893 Million cell updates/sec

Title: US-09-602-597-4

Perfect score: 1101
Sequence: 1 MAEPKSKVSTRTSSPAQGA.....GSSRYTLQASMGALNGVSP 216

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.67:*
2: PIR1:*
3: PIR2:*
4: PIR3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	741	67.3	157	A55046	plasmolipin - rat
2	158	14.4	281	T29464	hypothetical prote
3	153	13.9	153	A29472	T-cell surface gly
4	148.5	13.5	148	T38891	hypothetical prote
5	145.5	13.2	153	S68406	vesicular integral
6	141	12.8	156	T29152	hypothetical prote
7	95.5	8.7	242	T19664	hypothetical prote
8	95.5	8.7	444	T19294	McLeod syndrome-as
9	95	8.6	276	S13585	nosy protein precu
10	95	8.6	401	KOECNQ	sodium-glutamate
11	94.5	8.6	152	S32567	A4 protein - human
12	93	8.4	471	A81948	probable transmem
13	93	8.4	610	T32917	hypothetical prote
14	92	8.4	471	G81156	transporter, NacC
15	91	8.3	208	T23328	hypothetical prote
16	91	8.3	379	D81515	cell shape-determi
17	91	8.3	415	C72026	rod shape protein
18	90.5	8.2	404	LABECA	latent membrane pr
19	90	8.2	401	S42583	phosphate transloc
20	90	8.2	539	D69748	amino acid transpo
21	89.5	8.1	202	B83059	hypothetical prote
22	89	8.1	264	E69897	hypothetical prote
23	88.5	8.0	215	A84008	hemolysin III BH28
24	88.5	8.0	244	T22046	hypothetical prote
25	88.5	8.0	429	B75417	conserved hypotet
26	88.5	8.0	3005	T08841	polyprotein - dour
27	88	8.0	331	T23504	hypothetical prote
28	87.5	7.9	464	A75615	conserved hypotet
29	87	7.9	506	T29968	hypothetical prote

30	86.5	7.9	400	2	A41679	Beta-3-adrenergic
31	86.5	7.9	400	2	A53281	beta 3-adrenergic
32	86	7.8	406	2	T23292	hypothetical prote
33	85.5	7.8	519	2	T23739	hypothetical prote
34	85	7.7	358	2	D82989	alginate biosynthe
35	85	7.7	516	2	H70935	hypothetical prote
36	85	7.7	1147	2	A82340	sensor histidine k
37	84.5	7.7	314	2	E81064	tellurite resistan
38	84.5	7.7	397	2	E69094	conserved hypotet
39	84.5	7.7	414	2	S34829	triose phosphate/3
40	84	7.6	157	2	S62842	H+-transporting Ar
41	84	7.6	297	2	F64378	sodium-dependent p
42	84	7.6	315	2	T11388	NADH dehydrogenase
43	84	7.6	334	2	E71680	cytochrome C oxid
44	84	7.6	430	2	B82096	conserved hypotet
45	83.5	7.6	283	2	F75395	ABC transporter, p

ALIGNMENTS

```
RESULT 1
A55046
plasmolipin - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 18-Nov-1994 #sequence, revision 18-Nov-1994 #text_change 29-Sep-1999
C:Accession: A55046
R: Fischer, I., Sapirstein, V.S.
J. Biol. Chem. 269, 24912-24919, 1994
A:Title: Molecular cloning of plasmolipin. Characterization of a novel proteolipid re
A:Reference number: A55046, M01D:95014262
A:Accession: A55046
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-157 <FIS>
A:Cross-references: GB:U13617; NID:G532799; PIDN:AA62133.1; PID:G532800
C:Superfamily: T-cell surface glycoprotein MAL
C:Keywords: Lipoprotein; transmembrane protein

Query Match 67.3% Score 741; DB 2; Length 157;
Best Local Similarity 87.9%; Pred. No. 3.9e-56;
Matches 138; Conservative 13; Mismatches 6; Indels 0; Gaps 0;

QY 26 LRPDGFVRSRLGALMLQLVGLVMAIADPTPHLPAYGWMFVAFMLVTYIFN 85
      :|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 1 MRPDGFVRSALGVALLQLVGLVMAIADPTPHLPAYGWMFVAFMLVTYIFN 60
      :|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

QY 86 LYLQLAHMKLYMPWPLVMTFNISATVLYTFAIACSAVDLTLRGTREYNORAAAF 145
      :|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 61 IYLPQLHMKLYMPWPLVMTFNISATVLYTFAIACSAVDLTLRGTREYNORAAAF 120
      :|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

QY 146 FACLYMAYGVSAFFSYAMRGVGSNATSONAGGYA 182
      :|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 121 FACLYMAYGVSAFFSYAMRGVGSNATSONAGGYA 157
      :|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

RESULT 2
T29464
hypothetical protein F28H1.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence, revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T29464
R:Du, Z.; Le, T.T.
submitted to the EMBL Data Library, April 1997
A:Description: The sequence of C. elegans cosmid F28H1.
A:Reference number: Z20622
A:Accession: T29464
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-281 <DUZ>
A:Cross-references: EMBL:U97013; PIDN:AA52340.1; GSPDB:GN00019; CESP:F28H1.4
A:Experimental source: Strain Bristol NZ; Clone F28H1
```


DB 141 HAFSIV 146

RESULT 5

S68406

vesicular integral membrane protein VIP17 - dog

N:Alternate names: proteolipid VIP17; T-cell-specific MAL protein homolog VIP17

C:Species: Canis lupus familiaris (dog)

C>Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 29-Sep-1999

C:Accession: S68406; S68399

R:Zachetti, D.; Peranen, J.; Murata, M.; Fiedler, K.; Simons, K.

FEBS Lett. 377, 465-469, 1995

A:Title: VIP17/MAL, a proteolipid in apical transport vesicles.

A:Reference number: S68399; MUID:96140569

A:Accession: S68406

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-153 <ZAC>

A:Cross-references: EMBL:X92505; NID:q1054712; PIDN:CA63261.1; PID:q1054713

A:Accession: S68399

A:Molecule type: protein

A:Residues: 3-16;120-131 <ZAM>

A:Experimental source: kidney; cell-line Madin-Darby canine kidney cell

C:Function:

A:Pathway: myelin biogenesis

C:Superfamily: T-cell surface glycoprotein MAL

Query Match

Best Local Similarity 27.5%; Pred. No. 1.4e-05;

Matches 46; Conservative 25; Mismatches 73; Indels 23; Gaps 4;

DB 5

PKRVSTRTSSPAQGAASVARSALRPLDGFVRSRLGALMLQLVGLLWALADTPYHLYP 64

3 PAAASGSSSLP---SGFSVFTEFPDL-----LFEFFIRGGLWMIILASSLPVPL 50

DB 65

AAGWAFVAFVPLVTVIVFNLYLPLQLH-----MKIYMPWPLVLMIFNLSAVLYITAP 119

51 VGGWVAFVAFVCFEMATALLVLYIIAGHGENSWTLDAAHICIAALFTLASVLEALAP 110

DB 120

IACSAVNDLSLRGFRPYNQRAASFACTVMAVGVSAFVSQAWR 166

111 IGMQSEYTY-----KQYHENSIAVVFSTVATLLVYHAFSLIRMK 151

DB

RESULT 6

T29152

hypothetical protein F47B3.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T29152

R:Du, Z.; Le, T.T.

submitted to the EMBL Data Library, April 1997

A:Description: The sequence of C. elegans cosmid F47B3.

A:Reference number: 220579

A:Accession: T29152

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-156 <DUC>

A:Cross-references: EMBL:U97017; PIDN:AAB52358.1; GSPDB:GN00019; CESP:F47B3.3

A:Experimental source: strain Bristol N2; clone F47B3

C:Genetics:

A:Gene: CESP:F47B3.3

A:Map position: 1

A:Introns: 25/1; 63/3; 110/3

Query Match

Best Local Similarity 12.8%; Score 141; DB 2; Length 156;

Matches 45; Conservative 23; Mismatches 63; Indels 22; Gaps 5;

DB 40

LMHLDVGLLWALADTPYHLYPFAVPLVTVIVFNLYLPLQLH-----MKIYMPWPLVLMIFNLSAVLYITAP 99

DB 1 MKIVICVLCLEFIFVMMGPAY-YTGVNATFVSSVIGIFVTSLLTLVLFRRVDLPISIN 59

QY 100 WPLVY-----MINISATVIXITAFICSAADLTSLRGFRPNORAAASFAC 148

DB 60 WIVCSTRETEIIFQEMVYCFATVFFFLA--ACVLAVASOFRGTFAW--AIAAFAP 114

QY 149 LVMIAGVSAFVSQAWRC-----VGSNATSG 176

DB 115 GAAVCAFGPCVYKELKLSWKNNRATGGSNPVITQ 147

RESULT 7

T19664

hypothetical protein C33A12.9b - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000

C:Accession: T19664

R:Wilkinson, J.

submitted to the EMBL Data Library, January 1996

A:Reference number: Z19159

A:Accession: T19664

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-242 <WIL>

A:Cross-references: EMBL:268493; PIDN:CA92794.1; GSPDB:GN00022; CESP:C33A12.9b

A:Experimental source: clone C33A12

C:Genetics:

A:Gene: CESP:C33A12.9b

A:Map position: 4

A:Introns: 60/3; 113/3; 181/3

C:Superfamily: Caenorhabditis elegans hypothetical protein C33A12.9b

Query Match

Best Local Similarity 8.7%; Score 95.5; DB 2; Length 242;

Matches 32; Conservative 25; Mismatches 43; Indels 53; Gaps 5;

DB 42

LQVLYGLVWALADTPYHLYPAG-----WVWFV-----A 73

65 LIRVIVPLILFYPICT--FYLLIPALGVCRRPEYFGAWIYTTNSAGLRNSFNLYS 123

DB 74

VFLMIVTVIVFNLYL-----QLMKLYMPWPLVLMIFNLSATVLYITAFICSAADLT 129

124 IIFWLVASVIVNLFELVKKYQAKTOLVOIPQSYKAEFSITVTLVVIIFILNGC---- 179

DB

QY 130 SLRGFRPYNQRAASFACTVMAVGVSAFVS 162

180 -----FVLYMFVSYSYSPTSY 196

DB

RESULT 8

I39294

McLeod syndrome-associated protein XK - human

N:Alternate names: probable membrane transport protein

C:Species: Homo sapiens (man)

C>Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 01-Dec-2000

C:Accession: I39294; S69126

R:Ho, M.; Chelly, J.; Carter, N.; Dane, A.; Crocker, P.; Monaco, A.P.

Cell 77, 869-880, 1994

A:Title: Isolation of the gene for McLeod syndrome that encodes a novel membrane trar

A:Reference number: A54300; MUID:94273191

A:Accession: I39294

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-444 <REMS>

A:Cross-references: EMBL:232684; NID:9515872; PID:9515873

R:Rhamilich, S.; Bailey, P.; Blanchard, D.; Goossens, D.; Cartton, J.P.; Bertrand, O.

Eur. J. Biochem. 228, 931-934, 1995

A:Title: Purification and partial characterization of the erythrocyte Kx protein defi

A:Reference number: S69126; MUID:95255304

A:Accession: S69126

A:Status: preliminary

A:Molecule type: protein

A:Residues: 7-22 <KHA>
C:Genetics:
A:Gene: GDB:XK
A:Cross-references: GDB:120499; OMIM:314850
A:Map position: Xp21.2-Xp21.1
C:Keywords: phosphoprotein; transmembrane protein

Query Match	8.7%;	Score	95.5;	DB	2;	Length	444;
Best Local Similarity	22.4%;	Pred. No.	0.65;				
Matches	39;	Conservative	28;	Mismatches	66;	Indels	41;
				Gaps	8;		

```

QY 13 SSPGOGAEASYSARPDJGFEYSRGLMLOJVLGLLWALLA-DPBYLHYA-WGWM 70
      : : : : : : : : : : : : : : : : : : : : : : : :
Db 148 SAPOLITLOIYISWAGQDYVGRSLMTISLISLISYGCALRONILAIKKIDIEYKXNRRA 207
      : : : : : : : : : : : : : : : : : : : : : : : :
QY 71 FVAVFLW-----LVYIVLENLYLEFOLHMKLYMWBPVLMIFNISATVLYTAFIAC 122
      : : : : : : : : : : : : : : : : : : : : : : : :
Db 208 YVCJEFIMRSEPIATRRVYVLYLFISVL-----KMWVYIILINFESFLYKPLLFWC 258
      : : : : : : : : : : : : : : : : : : : : : : : :
QY 123 SAAVDLSLNGTRPYNO-----RAAASEFACLVMIAY-GVSAFFSQCAMRGV 168
      : : : : : : : : : : : : : : : : : : : : : : : :
Db 259 SGS-----PPEENIEKALSRVOTIYLCVLTILLYIGIMF-----CWSAV 298
      : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT

nosty protein precursor - Pseudomonas stutzeri
 C:Species: Pseudomonas stutzeri
 C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 08-Oct-1995
 C:Accession: S13585
 R:Zunf, W.G., Viedrock-Sambale, A.; Braun, C.
 Eur. J. Biochem. 192, 591-599, 1990
 A:Title: Nitrous oxide reductase from denitrifying Pseudomonas stutzeri. Genes for copper
 proteins.
 A:Reference number: S13582; MUID:91006150
 A:Accession: S13585
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-276 <ZUN>
 A:Cross-references: EMBL:X53676; NID:g45844; PIDN:CAA37717.1; PID:g45850

Query Match	8.63;	Score 95;	DR 2;	Length 276;
Best Local Similarity	24.38;	Pred. NO. 0.46;		
Matches 43;	Conservative 31;	Mismatches 55;	Indels 48;	Gaps 9;

QY 23 VSALRPDLGFRSLGALMLDLY-IGLLVMA-----LIDTPPHKLPAAGGWVF----- 71
 Db 111 ILALVILIGGCALALIALIVEGVELGMLFMAWGKRMISST-----LIGVFLAAYVL 164
 QY 72 -----VAFILMLVILENLYEQL-----HKKLMYBPMPLIMFNISATY 113
 Db 165 SGKVNKSSAAGLALGWLFYLVLFDFDLVLLALVLSBKGNPELLPMLLL-----NPTD 219
 QY 114 LY-----TFAIACSAVDLTLSTGRTRPYNORAASFPACILMAYGSAFESYQAMR 166
 Db 220 IYRLINSGFEGSSAGAGVSLGADLPV---PAAVIMLCI-LAMIGVSLILAAIAIR 272

RESULT 10

sodium-glutamate symport carrier protein - *Escherichia coli*.
 N:Alternate names: glutamate permease
 C:Species: *Escherichia coli*
 C:Date: 30-Jun-1991 #sequence_revision 31-Oct-1997 #text_change 16-Jun-2000
 C:Accession: G65166; A36524; S14031; J00489
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co-
 A.: Rose, D.J.; Mau, B.; Shaq, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of *Escherichia coli* K-12.
 A:Reference number: A64720; M01D:97426617
 A:Accession: G65166

A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1401 nt

A:Residues: 1-401 <BAT>
A:Cross-references: GB: E000442; GB: U00096; NID: g2367253; PIDN: AAC76677.1; PID: g179000
A:Experimental source: strain K-12, substrain MG1655
R:Deguchi, Y.; Yamato, I.; Anraku, Y.
J. Biol. Chem. 265, 21704-21708, 1990
A:Title: Nucleotide sequence of *gltS*, the Na⁺/glutamate symport carrier gene of *Escherichia coli* K-12
A:Reference number: A36524, MIMD: g1072371

A: molecule type: DNA

A:Residues: 1-377, 'S', 379-401 <DEG>
A:Cross-references: GB:D00626; NID:g216540; PIDN:BAA00517.1; PID:g216541

A; Experimental source: strain B29-78
R; Kalman, M.; Gentry, D.R.; Cashel, M

Mol. gen. genet. 423, 379-386, 1991
 A1: Title: Characterization of the Escherichia coli K12 glts glutamate permease gene
 A: Reference number: S14031, MUID:91203811

A;Molecule type: DNA

A/Residues: 1-401 <NAL>
A/Cross-references: EMBL:X17499; NID:g41592; PIDN:CAA35540.1; PID:g415933
A/Contact:

A;Gene: g1t

A:map position: 82 min
 C:Superfamily: sodium-glutamate symport carrier protein
 C:keywords: glutamate transport; transmembrane protein

Query Match	8.6%	Score 95	DB 1	Length 401
Best Local Similarity	22.6%	Pred. NO	0.65	
Matches 47	Conservative 35	Mismatches 96	Indels 30	Gaps 6

```

0Y      1 MAEPPSKVSTRTSSPAQGAESVALREDLGEVBSRLG-----ALMLQLVYLGLLWALI 55
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     179 VARLVKHSHTPNGLPDDQEVPTAFERKDVGRMTLSLVLETIALIAICLVGKIVAO LL 238

```

```

Oy 56 ADPPYHLPRAGWMFVAFLEMLVTIVLEFNLYLEQLHMKLYMPWMBLVIAMFINISAFVLY 115
      | | : | . | : | : | | | : | | | | : | : | : | : | : | : | : | : |
Db 239 AGTAFEL-PFVVCVLEFEGVILISGLST-----MGRYRVEFRAVSVLGNSLSLEL 287

```

QY 116 ITAFIACSAVDLTSLRTPRYNQRAASFPACLMIATGVSAFESYQAWRGVGSNAATS 175
+ + + + +
Db 288 AMALMGK-LMELAST-----ALPMLAIIVQTITMALYAIFVTWRMGKNDA 336

```
OY      176 QMAGGYATCATAPSGAEAAAGSQSRVT 2033
          :| | : | :
Db      337 VLAAGH---CGFGLGATPTAIANMOAIT 3611
```

RESULT 11

A4: protein - human
C: Species: Homo sapiens (man)
C: Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000
C: Accession: S32567
R: Oliva, M.M.; Wu, T.C.; Yang, V.W.
Arch. Biochem. Biophys. 302, 183-192, 1993
A: Title: Isolation and characterization of a differentiation-dependent gene in the hu
C: Reference number: S32567; M01D: S328341

A;Status: preliminary

A;Molecule type: mRNA
A;Residues: 1-152 <OLI>

A/Cross References: EMDL: P03604; NID: g1/1893; PIDN: AAA33433.1; PID: g1/1500
C/Genetics:
A/Gene: GDB:M

A; Cross-references: GDB:683190

Query Match	8.6%	Score 94.5	DB 2	Length 152
Best Local Similarity	21.3%	Pred. No. 0.3		
Matches 38; Conservative	32	Mismatches	71	Indels 37; Gaps 7

RESULT 15

T23328

hypothetical protein K04G2.9 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T23328

R:Gardner, A.

submitted to the EMBL Data Library, July 1996

A:Reference number: 219727

A:Accession: T23328

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-208 <M11>

A:Cross-references: EMBL:275712; PIDN:CAB00046.1; GSPDB:GN00019; CESP:K04G2.9

A:Experimental source: clone K04G2

C:Genetics:

A:Gene: CESP:K04G2.9

A:Map position: 1

A:Introns: 22/2; 57/3; 175/3

Query Match

Best Local Similarity 8.3%; Score 91; DB 2; Length 208;

Matches 45; Conservative 30; Mismatches 76; Indels 32; Gaps 8;

```

QY 8 VSTRSSPAQGAASVSL--RPDLGFVRSRLGALMLQVLGLLVWALADPPYHLP 64
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 17 VTTTREKTKYKEHSWSCGHGPTDKNYCIGPLRIAEIPLCVIVCLITSV-FGPGP 75
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 65 AYGWVWF-----VAVFIMLVITVLENLXLEQLMK-LYMPMPVLMIPINISATVLYI 116
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 76 FRG-ILFEGTILLTVASVAMLTFFFLAYFFTLHLSHDFECWREADLLENVICAVLEFI 134
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 117 TAFIA-----CSAVIDLTS---LRGTRPYNOBAASPFACIYMIAYGVSAF 160
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 135 VLSIYEAAYSTGSMNCHN---DISGDLINHGCRITTEMASFILFTLGIFALTAPL 191
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 161 STQ 163
   | | |
DB 192 THK 194

```

Search completed: May 2, 2001, 15:11:00
 Job time: 65 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2001, 15:10:20 / Search time 9.56 Seconds

(Without alignments)
773.973 Million cell updates/sec

Title: US-09-602-597-4

Perfect score: 1101
1 MAEPSSKYSTRSSPAQGA.....GSQSRVTLQASMGALRGVSP 216

Sequence: BLOSOM62
Gapop 10.0, Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	930	84.5	182	1 PLIP_HUMAN	O9Y342 homo sapien
2	848	77.0	182	1 PLIP_RAT	P47987 rattus norv
3	155	14.1	153	1 MAL_MOUSE	O09198 mus musculu
4	153	13.9	153	1 MAL_HUMAN	P21145 homo sapien
5	152	13.8	153	1 MAL_RAT	O64349 rattus norv
6	148.5	13.5	148	1 BENE_HUMAN	Q13021 homo sapien
7	145.5	13.2	153	1 MAL_CANFA	O28296 canis famli
8	131	11.9	296	1 MYDM_MOUSE	O35682 mus musculu
9	95	8.6	276	1 NOSY_PSEST	P19845 pseudomona
10	95	8.6	401	1 GUTS_ECOLI	P19933 escherichia
11	94.5	8.6	152	1 AAP_HUMAN	O04941 homo sapien
12	93.5	8.5	444	1 AAP_HUMAN	P51811 homo sapien
13	92	8.4	496	1 SMVA_SALTY	P37594 salmonella
14	90.5	8.2	404	1 LMPL_EBYC	P28362 epstein-bar
15	86.5	7.9	400	1 B3AR_RAT	P26255 rattus norv
16	86	7.8	351	1 B3AR_CAVPO	O60483 cavia porce
17	84	7.6	157	1 Y406_MYCPN	O50325 mycoplasma
18	84	7.6	297	1 Y630_METJA	O58047 methanococc
19	82	7.4	455	1 ENTL_HUMAN	O99808 homo sapien
20	81.5	7.4	400	1 B3AR_MOUSE	P25962 mus musculu
21	81	7.4	528	1 YEPD_YEAST	P40039 saccharomyc
22	80.5	7.3	428	1 GPRX_ORILA	O91178 oryzias lat
23	80	7.3	305	1 NDIW_APLTA	P34847 apis mellif
24	80	7.2	394	1 CC91_YEAST	P41733 saccharomyc
25	79.5	7.2	405	1 B3AR_CANFA	O02662 canis famli
26	79.5	7.2	476	1 GREI_BALAM	O93126 balanus amp
27	79.5	7.2	555	1 CSTR_AQUAE	O67304 aquilex aeo
28	79	7.2	405	1 B3AR_CAPHI	O9457 capra hircu
29	78	7.1	678	1 YF48_MYCTU	O10778 mycobacteri
30	77.5	7.0	379	1 CYB_SORCI	O79451 sorcx ciner
31	77.5	7.0	379	1 CYB_SORCI	O79464 sorcx tundi
32	77	7.0	312	1 ORF4_RAT	P23273 rattus norv
33	77	7.0	318	1 YBHN_ECOLI	P75770 escherichia

34	77	7.0	412	1 Y653_METTH	O26759 methanobact
35	77	7.0	428	1 BCH2_RHOCA	P26171 rhodobacter
36	77	7.0	468	1 B1AR_PIG	O28998 sus scrofa
37	77	7.0	525	1 DIE2_YEAST	P50076 saccharomyc
38	76.5	6.9	152	1 TOX4_BORPE	P04960 bordecella
39	76.5	6.9	318	1 NDIW_TRIVU	O78707 trichosurus
40	76.5	6.9	355	1 OPSB_ASTFA	P51472 astyanax fa
41	76.5	6.9	395	1 Y306_MYCPN	P75343 mycoplasma
42	76.5	6.9	401	1 YEDE_SALTY	O06400 salmonella
43	76.5	6.9	546	1 LNT_TREPA	O83279 treponema p
44	76	6.9	226	1 ATP6_MACRO	P92664 macropus r
45	76	6.9	463	1 Y096_MYCTU	O10892 mycobacteri

ALIGNMENTS

RESULT 1
ID PLIP_HUMAN STANDARD; PRT; 182 AA.
AC O9Y342; 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
OS PLASMOLOPIN.
OC Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RA Xie B., Durrle R., Sapirstein V.S.;
RT Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: APPEARS TO BE INVOLVED IN MYELINATION. COULD ALSO PARTICIPATE IN ION TRANSPORT EVENTS AS ADDITION OF PLASMOLOPIN TO LIPID BILAYERS INDICES THE FORMATION OF ION CHANNELS, WHICH ARE VOLTAGE-DEPENDENT AND K(+)-SELECTIVE (BY SIMILARITY).
CC -1- SUBUNIT: HEXAMER ARRANGED AS A TRIMER OF TWO PLASMOLOPIN SUBUNITS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE MAL FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
DR EMBL; AF137386; AAD33060.1; -
DR MIM; 600340; -
KM Transmembrane; Transport; Ionic channel; Ion transport.
FT DOMAIN 1 35 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 36 56 POTENTIAL.
FT DOMAIN 57 68 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 69 89 POTENTIAL.
FT DOMAIN 90 99 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 100 120 POTENTIAL.
FT DOMAIN 121 141 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 142 162 POTENTIAL.
FT DOMAIN 163 182 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 182 AA; 19986 MW; 6509AA080E267E8 CRC64;

Query Match 84.5%; Score 930; DB 1; Length 182;
Best Local Similarity 100.0%; Pred. No. 9.6e-63;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MAEPSSKYSTRSSPAQGAESVSLRPDLGPFVSRRLNKLQVLGLWALIAIDPPY 60
|||||

RESULT 5
MAL_RAT STANDARD; PRT; 153 AA.
AC 064349;
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE MYELIN AND LYMPHOCYTE PROTEIN (T-LYMPHOCYTE MATURATION-ASSOCIATED
DE PROTEIN) (17 KDA MYELIN VESICULAR PROTEIN) (MVP17) (NS 3).
GN MAL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Brain;
RA MEDLINE=96129665; PubMed=9583510;
RA Kim T., Fiedler K., Madison D.L., Krueger W.H., Pfeiffer S.E.;
RT "Cloning and characterization of MVP17: a developmentally regulated
RT myelin protein in oligodendrocytes";
RL J. Neurosci. Res. 42:413-422(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-LEWIS; TISSUE=Spinal cord;
RA MEDLINE=95370938; PubMed=7643216;
RA Scharen-Wiemers N., Valenzuela D.M., Frank M., Schwab M.E.;
RT "Characterization of a rat gene, RML, encoding a protein with four
RT hydrophobic domains in central and peripheral myelin";
RL J. Neurosci. 15:5753-5764(1995).
CC -1- FUNCTION: COULD BE AN IMPORTANT COMPONENT IN VESICULAR TRAFFICKING
CC CYCLING BETWEEN THE GOLGI COMPLEX AND THE APICAL PLASMA MEMBRANE.
CC COULD BE INVOLVED IN MYELIN BIOGENESIS AND/OR MYELIN FUNCTION.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: EXCLUSIVELY EXPRESSED IN WHITE AND GRAY MATTER
CC OLIGODENDROCYTES IN THE CNS AND IN MYELINATING SCHWANN CELLS IN
CC PERIPHERAL NERVES. HIGHER LEVELS ARE FOUND IN THE WHITE. WEAK
CC CNS OR SPINAL CORD. IN THE GRAY MATTER THAN IN THE WHITE.
CC EXPRESSION ALSO FOUND IN SPLEEN AND KIDNEY.
CC -1- DEVELOPMENTAL STAGE: DETECTED JUST BEFORE BIRTH IN SCHWANN CELLS
CC AND AFTER BIRTH IN OLIGODENDROCYTES OF BRAINSTEM AND SPINAL CORD.
CC -1- INDUCTION: UPREGULATED DURING MYELINATION.
CC -1- PM: LIPOPROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE MAL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U13367; AAC53366.1; -
DR EMBL; X82557; CA837903.1; -
FT Transmembrane; Lipoprotein; Myelin.
FT DOMAIN 1 24
FT TRANSSEM 25 46
FT TRANSSEM 47 53
FT TRANSSEM 54 75
FT TRANSSEM 76 92
FT TRANSSEM 93 114
FT TRANSSEM 115 125
FT TRANSSEM 126 147
FT TRANSSEM 148 153
FT DOMAIN 153 153
FT SEQUENCE 153 AA; 16758 MW; C98996DC3612494 CRC64;
Query Match 13.8%; Score 152; DB 1; Length 153;
Best Local Similarity 29.9%; Pred. No. 4.6e-05;
Matches 44; Conservative 23; Mismatches 66; Indels 14; Gaps 4;

QY 22 SVSALRPDLGVRSRRLGMLQVLVALLADPPIHXPAYGWMFVAPLMLVTI 81
DB 17 SVFTFPDL-----LFFEFIFGGLVWILLASSLPVLOGWMEFVSCFLATT 67
QY 82 VLENLYEQLHMKLYMPWPLVLMIFNSATVLTAFIACSAVDLSLNG--TRPNO 139
DB 68 SLAWMTIIGHG--ETSMITLDAHYCAALFTLSASV-LEALATTMFDGFTYRHYE 124
QY 140 RAASFPACLVMIYGVSAFESYQAWR 166
DB 125 NIAAVFAVATILLYHAFVSLRMK 151
RESULT 6
BENE_HUMAN STANDARD; PRT; 148 AA.
AC 013021;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE BENE PROTEIN (FRAGMENT).
GN BENE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95347597; PubMed=7622049;
RA Lautner-Rieske A., Thiede R., Zachau H.G.;
RT "Searching for non-v kappa transcripts from the human immunoglobulin
RT kappa locus";
RL Gene 159:199-202(1995).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE MAL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U17077; AAA76738.1; -
DR KIM; 602222; -
FT Transmembrane 1
FT NON_TER 1
FT TRANSSEM 17 37
FT TRANSSEM 34 74
FT TRANSSEM 92 112
FT TRANSSEM 126 146
FT SEQUENCE 148 AA; 16849 MW; 561955796EE78DAB CRC64;
Query Match 13.5%; Score 148.5; DB 1; Length 148;
Best Local Similarity 30.2%; Pred. No. 8.1e-05;
Matches 38; Conservative 31; Mismatches 52; Indels 5; Gaps 4;
QY 39 ALMLDQVLDLVALLADPPIHXP-AYGVVAVVAVFLVTVLPVLFQVLMKLYM 97
DB 24 AFLPELIFGLFWTVAAT-HIYPLDLOGVVMVSLTSLIMFLSLYFGFYKR--F 80
QY 98 VPMPLVIMNISATVLYIT-AFIACSAVDLTSLGTRPYNOBAAAFACLVMIYGV 156
DB 81 ESRVADSLVHGTLGIIYMAAVYQVATVSEKLDPRITTYINSASFATITLITIL 140
QY 157 SAFFSY 162
DB 141 HAFSIV 146

```

RESULT 7
MAL_CANFA STANDARD; PRT; 153 AA.
ID MAL_CANFA
AC Q28296;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE MELIN AND LYMPHOCYTE PROTEIN (T-LYMPHOCYTE MATURATION-ASSOCIATED
DE PROTEIN) (VIP17 PROTEOLIPID).
GN MAL.
OS Caris familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=96140569; PubMed=8549777;
RA Zaccaretti D., Peranen J., Murata M., Fiedler K., Simons K.;
RT "VIP17/MAL, a proteolipid in apical transport vesicles.";
RL FEBS Lett. 377:465-469(1995).
CC -1- FUNCTION: COULD BE AN IMPORTANT COMPONENT IN VESICULAR TRAFFICKING
CC CYCLING BETWEEN THE GOLGI COMPLEX AND THE APICAL PLASMA MEMBRANE.
CC COULD BE INVOLVED IN MYELIN BIOGENESIS AND/OR MYELIN FUNCTION.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE MAL FAMILY.
CC -1- PFM: LIPOPROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X92505; CA63261.1;
DR Transmembrane; 1; Lipoprotein; Myelin.
KW DOMAIN 1 24 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 25 46 POTENTIAL.
FT DOMAIN 47 53 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 54 75 POTENTIAL.
FT DOMAIN 76 92 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 93 114 POTENTIAL.
FT DOMAIN 115 125 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 126 147 POTENTIAL.
FT DOMAIN 148 153 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 153 AA; 16624 MW; 3E521B023B9CDB9B CRC64;

Query Match 13.2%; Score 145.5; DB 1; Length 153;
Best Local Similarity 27.5%; Pred. No. 0.00014;
Matches 46; Conservative 25; Mismatches 73; Indels 23; Gaps 4;

OY 5 PSKSTRSSPAQGAASVSLRPDLGFVRSRLGALMLQVLGLVWALIDTPYHLYP 64
DB 3 PAAAGSGGSLP---SGEFTFFPDLP-----LTFEEFFGGLWILLASSIVPILP 50
OY 65 AYGWMEFAVFLMVTIYVLFNXLPLH-----MKLWMPWPLVLMIFMISATVLYTAF 119
DB 51 VQGWMEFVSFCEMATTALVLTITGAGGENSWTLDAAHYCAALFTLSASVLEALMT 110
OY 120 IACSAVDLTLKRTPRYNQRAAASFACLVMIAGVSAFESYQAMR 166
DB 111 IGMGSGTY-----KQYHENISAVFVSVALLYVHAHFSLIRK 151

```

```

DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MYELOID-ASSOCIATED DIFFERENTIATION MARKER (MYELOID UPREGULATED
DE PROTEIN).
GN MYADM OR MOG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Scurionathli; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Pettersson M., Nilsson K., Jonsson J.I.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE MAL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AJ001616; CA04870.1;
DR MGD; MGI:1355332; Myadm.
KW Transmembrane; 35 55 POTENTIAL.
FT TRANSSEM 61 81 POTENTIAL.
FT TRANSSEM 95 115 POTENTIAL.
FT TRANSSEM 131 151 POTENTIAL.
FT TRANSSEM 166 186 POTENTIAL.
FT TRANSSEM 197 217 POTENTIAL.
FT TRANSSEM 225 245 POTENTIAL.
FT TRANSSEM 263 283 POTENTIAL.
SQ SEQUENCE 296 AA; 32098 MW; 5BD37C839734AD9D CRC64;

Query Match 11.9%; Score 131; DB 1; Length 296;
Best Local Similarity 25.4%; Pred. No. 0.0029;
Matches 45; Conservative 26; Mismatches 76; Indels 28; Gaps 6;

OY 8 VSTRSSPAQGAASVSLRPDLGFVRSRLGALMLQVLGLVWALIDTPYHLYP 67
DB 10 ITTTSSST--TVGSARALTPGLR-----LLQLSICVAFSLVSAVGARTGPGN 60
OY 68 WYMEFAVFLMVTIYVLFNXLPLH-----FQLEHKLWMPWPLVLT---MFINASVLYTAF 121
DB 61 WAMFTWCFENATLILLYELGGLAHPLSRNRPPIFACVAAALFCLSSSIYPTTV- 119
OY 122 CSAVDLTLKRTPRYNQRAAASFACLVMIAGVSAFESYQAMR 178
DB 120 -----OFLAAGRTRDAIATTFSCVACLAIVATEV-----AMTRAPGEITGYMA 164

RESULT 9
NOSTY_PSEST STANDARD; PRT; 276 AA.
ID NOSTY_PSEST
AC P19845;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 01-MAR-1992 (Rel. 21, Last annotation update)
DE MEMBRANE PROTEIN NOSTY PRECURSOR.
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=316;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 14405 / ZOBELL;
RX MEDLINE=91006150; PubMed=2170125;
RA Zunft W.G., Viebrock-Sambale A., Braun C.;
RT "Nitrous oxide reductase from denitrifying Pseudomonas stutzeri.
RT Genes for copper-processing and properties of the deduced products,

```

RT Including a new member of the family of ATP/GTP-binding proteins.",
 RL Eur. J. Biochem. 193:591-599(1990).
 CC -1- FUNCTION: INVOLVED IN COPPER PROCESSING.
 CC -1- SUBCELLULAR LOCATION: PLASMA MEMBRANE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X53676; CAA37717.1; -
 DR PIR: S13585; S13585.
 KM Transmembrane; Copper; Signal.
 FT SIGNAL 1 34
 FT CHAIN 35 276 MEMBRANE PROTEIN NOSY.
 SQ SEQUENCE 276 AA; 29434 MW; 91411E4F0AD8D15 CRC64;

Query Match 8.6%; Score 95; DB 1; Length 276;
 Best Local Similarity 24.3%; Pred. No. 1.3;
 Matches 43; Conservative 31; Mismatches 55; Indels 48; Gaps 9;

QY 23 VSALRPDGFVRSRLGALMLQLV-LGLVVA-----LADTPHYLYRAYGVNF----- 71
 DB 111 IIALVLVLFQFCALALALVEGVGLGLFVAFGRFMIST-----LLGVVLAFAVYL 164
 QY 72 -----VAFLVLTIVLENTLFLQ-----HMLVYVMPVLMIFINISATV 113
 DB 165 SGKVKESKSSAGLALGVFLVFDVLLVLLVLSSEKFPPELLPMLL-----NPTD 219
 QY 114 LY-----ITAFIACSAANDLTSLGRPYNORAAAFACLVMIAYGVSAFESYQAMR 166
 DB 220 IYRLINLSGFESGSSAMGVLSIGADLPV---PAAVLMCL-LAMIGVSLIAYIAIFR 272

RESULT 10
 GLTS_ECOLI STANDARD; PRT; 401 AA.
 AC P19933;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE SODIUM/GLUTAMATE SYMPORT CARRIER PROTEIN (GLUTAMATE PERMEASE).
 GN GLTS OR GLTC.
 OS Escherichia coli.
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia.
 OX NCBI_TaxID=562;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-B:
 RC MEDLINE-91072371; PubMed-2254324;
 RA Deguchi Y., Yamato I., Anraku Y.,
 RT "Nucleotide sequence of glts, the Na⁺/glutamate symport carrier gene
 RL of Escherichia coli B.",
 RL J. Biol. Chem. 265:21704-21708(1990).
 RP [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN-K12;
 RC MEDLINE-91203811; PubMed-2017136;
 RA Kalman M., Gentry D., Cashel M.,
 RT "Characterization of the Escherichia coli K12 glts glutamate permease
 RL gene".
 RL Mol. Gen. Genet. 225:379-386(1991).
 RP [3]
 RN SEQUENCE FROM N.A.
 RP STRAIN-K12 / MG1655;
 RC MEDLINE-93315143; PubMed-7686882;
 RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.,
 RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli

RT genome: organizational symmetry around the origin of replication.",
 RL Genomics 16:551-561(1993).
 RL [4]
 RP SEQUENCE OF 379-401 FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE-92184121; PubMed-1544582;
 RA Kalman M., Murphy H., Cashel M.,
 RT "The nucleotide sequence of recg, the distal spo operon gene in
 RL Escherichia coli K-12.",
 RL Gene 110:95-99(1992).
 CC -1- FUNCTION: CATALYZES THE SODIUM-DEPENDENT UPTAKE OF EXTRACELLULAR
 CC GLUTAMATE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D00626; BAA00517.1; -
 DR EMBL: X17499; CAA35540.1; -
 DR EMBL: L10328; AAA62006.1; -
 DR EMBL: AE000442; AAC76677.1; -
 DR EMBL: M6367; AAA28514.1; -
 DR PIR: J00489; YOECONQ.
 DR PIR: A36524; A36524.
 DR PIR: S14031; S14031.
 DR Ecogen: Egi0406; glts.
 KM Transport; Transmembrane; Inner membrane; Sodium transport; Symport.
 FT TRANSMEM 3 21 POTENTIAL.
 FT TRANSMEM 36 54 POTENTIAL.
 FT TRANSMEM 69 87 POTENTIAL.
 FT TRANSMEM 91 109 POTENTIAL.
 FT TRANSMEM 112 132 POTENTIAL.
 FT TRANSMEM 160 178 POTENTIAL.
 FT TRANSMEM 214 240 POTENTIAL.
 FT TRANSMEM 246 267 POTENTIAL.
 FT TRANSMEM 276 294 POTENTIAL.
 FT TRANSMEM 301 325 POTENTIAL.
 FT TRANSMEM 336 357 POTENTIAL.
 FT TRANSMEM 371 398 POTENTIAL.
 FT TRANSMEM 378 378 G -> S (IN REF. 1).
 FT CONFLICT 378 378
 SQ SEQUENCE 401 AA; 42425 MW; 9569AE0AF7634DC CRC64;

Query Match 8.6%; Score 95; DB 1; Length 401;
 Best Local Similarity 22.6%; Pred. No. 1.7;
 Matches 47; Conservative 35; Mismatches 96; Indels 30; Gaps 6;

QY 1 MAEPKSVTRTSSPAGAEVSALRPDGFVSRG-----ALMLQLVGLVVALI 55
 DB 179 VAVLVHSTTPNGIPDDQEVPAFEKPDVGRMITSVLLETIALICLVGVIVQL 238
 QY 56 ADPEHYLYPAGVWVFAVFLMTIVLENTLFLQ-----HMLVYVMPVLMIFINISATV 115
 DB 239 AGTAFEL-PFVCLVLFGLVLSNGLS-----MGFVRYERAVSVYGNLSLFL 287
 QY 116 ITAFIACSAANDLTSLGRPYNORAAAFACLVMIAYGVSAFESYQAMR 175
 DB 288 AMLMGLK-LWEIASEL-----ALPMLAIIVQITIFALYAIPTVBMGKNYDAA 336
 QY 176 QMAGVATTCATAPSGAAGSRTV 203
 DB 337 VLAAGH---CGFGLGATPTAIANQAIT 361

RESULT 11
 A4P_HUMAN STANDARD; PRT; 152 AA.
 ID A4P_HUMAN
 AC Q04941;

01-OCT-1993 (Rel. 27, Created)
01-OCT-1993 (Rel. 27, Last sequence update)
01-OCT-2000 (Rel. 40, Last annotation update)
INTESTINAL MEMBRANE A4 PROTEIN (DIFFERENTIATION-DEPENDENT PROTEIN A4)
(PROTEOLIPID PROTEIN 2).
PLP2 OR A4.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
MEDLINE=93228341; PubMed=8470895;
Olive M.M., Wu T.-C., Yang V.W.;
"Isolation and characterization of a differentiation-dependent gene
in the human colonic cell line HT29-18".
Arch. Biochem. Biophys. 302:183-192(1993).
[2]
SEQUENCE FROM N.A.
MEDLINE=9608923; PubMed=9344658;
Fisher S.E., Ciccodicola A., Tanaka K., Curci A., Deslato S.,
D'Urso M., Craig I.W.;
"Sequence-based exon prediction around the synaptophysin locus
reveals a gene-rich area containing novel genes in human proximal
XP".
Genomics 45:340-347(1997).
-1- FUNCTION: MAY PLAY A ROLE IN CELL DIFFERENTIATION IN THE
INTESTINAL EPITHELIUM.
-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
-1- TISSUE SPECIFICITY: ENRICHED IN COLONIC MUCOSA. THE EXPRESSION
OF A4 FOLLOWS A GRADIENT ALONG THE CRYPTO-VILLUS AXIS WITH THE
MOST ABUNDANT MESSAGE OCCURRING IN THE LOWER HALF OF THE CRYPT.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

DR EMBL; L09604; AAA5499.1; -
DR EMBL; U93305; AAB92356.1; -
DR PIR; S32567; S32567.
DR MIM; 300112; -
KM Transmembrane; glycoprotein.
FT TRANSMEM 25 45 POTENTIAL.
FT TRANSMEM 48 68 POTENTIAL.
FT TRANSMEM 85 105 POTENTIAL.
FT TRANSMEM 112 132 POTENTIAL.
FT CARBOHYD 18 18 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 108 108 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 152 AA; 16690 MW; 689378A8C326206C CRC64;

Query Match 8.6%; Score 94.5; DB 1; Length 152;
Best Local Similarity 21.3%; Pred. No. 0.84;
Matches 38; Conservative 32; Mismatches 71; Indels 37; Gaps 7;

QY 9 SRTSSPAQGAASVSALRPDLGFRSRGALMLQLVGLVW-ALINDPYHLYPAYG 67
DB 4 SERLSAPGCAACT-----NFSRIKGIILFAEIIICLVILICFSASTGY----- 49
QY 68 WYFAVFLMTVILFNLVLFQLMKLYVMPVPLVLMFNISAVLYTATACSAVD 127
DB 50 -SSLSVIEMLIAIFVYVYVMDLHTKIPFIMWPSDFRRLIALILV----- 96
QY 128 LLSL-----KGRPYNORAAAFACLVMIANGVSASFYQMRGVGSAATASQAGG 180
DB 97 TTSIVLVVERGN--HSKIVAGVGLIATGCLGDAYVTFPV-RQPRHTAAPTDPADG 150
RESULT 12

01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
01-OCT-2000 (Rel. 40, Last annotation update)
MEMBRANE TRANSPORT PROTEIN XK (XK ANTIGEN).
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
MEDLINE=94273191; PubMed=8004674;
Ho M., Chelly J., Carter N., Danek A., Crocker P., Monaco A.P.;
"Isolation of the gene for McLeod syndrome that encodes a novel
membrane transport protein".
Cell 77:869-880(1994).
[2]
REVISED TO 204-205.
HO M.F.;
Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
-1- FUNCTION: MAY BE INVOLVED IN SODIUM-DEPENDENT TRANSPORT OF NEUTRAL
AMINO ACIDS OR OLIGOPEPTIDES.
-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
-1- TISSUE SPECIFICITY: HIGH LEVELS IN SKELETAL MUSCLE, HEART, BRAIN,
AND PANCREAS; LOW LEVELS IN PLACENTA, LUNG, LIVER, AND KIDNEY.
-1- POLYMORPHISM: XK IS RESPONSIBLE FOR THE XK BLOOD GROUP SYSTEM.
-1- DISEASE: DEFECTS IN XK ARE THE CAUSE OF MCLEOD SYNDROME, AN X-
LINKED MULTISYSTEM DISORDER CHARACTERIZED BY ABNORMALITIES IN THE
NEUROMUSCULAR AND HEMATOPOIETIC SYSTEMS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

DR EMBL; 232684; CAA83632.1; -
DR MIM; 314850; -
KM Transmembrane; Transport; Amino-acid transport; Blood group antigen.
FT TRANSMEM 1 2 POTENTIAL.
FT TRANSMEM 3 23 POTENTIAL.
FT TRANSMEM 24 37 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 38 58 POTENTIAL.
FT TRANSMEM 59 68 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 69 89 POTENTIAL.
FT TRANSMEM 90 140 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 141 161 POTENTIAL.
FT TRANSMEM 162 171 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 172 192 POTENTIAL.
FT TRANSMEM 193 208 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 209 229 POTENTIAL.
FT TRANSMEM 230 235 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 236 256 POTENTIAL.
FT TRANSMEM 257 277 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 278 298 POTENTIAL.
FT TRANSMEM 299 317 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 318 338 POTENTIAL.
FT TRANSMEM 339 349 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 350 370 POTENTIAL.
FT TRANSMEM 371 444 CYTOPLASMIC (POTENTIAL).
SO SEQUENCE 444 AA; 50876 MW; E94BDD0E3BEF7AB2 CRC64;

Query Match 8.5%; Score 93.5; DB 1; Length 444;
Best Local Similarity 22.4%; Pred. No. 2.4;
Matches 39; Conservative 28; Mismatches 66; Indels 41; Gaps 8;

QY 13 SSPAGAAASVSALRPDLGFRSRGALMLQLVGLVWMLIA-DPYHLYPA-YGVW 70


```

Db      148  SAPOFLDLYISVMQODYVGRSLMTJTSLSISIVYGALRCNITLAIKIKDEYEVKALA 207
Oy      71  FAVALM-----LVYIVFNLYLFQHLKLVNPPVLMINISATVYITAFINC 122
Db      208  IYCIPLMSFEIATRVVVLFTSVL-----KTVWVYIILLNFSEFLYIPILFWC 258
Oy      123  SAVALDTSLRGTRPYNO-----RAASFACLVMIAV-GVSAFEQWAGV 168
Db      259  SGS-----PPENIERKLSRVGTTIVLCFLLTYTGIMF---CWSAV 298

RESULT 13
SMVA_SALTY
ID      SMVA_SALTY      STANDARD:      PRT:      496 AA.
AC      P37594;
DT      01-OCT-1994 (Rel. 30, Created)
DT      01-OCT-1994 (Rel. 30, Last sequence update)
DT      01-FEB-1995 (Rel. 31, Last annotation update)
DE      METHYL VIologen RESISTANCE PROTEIN SMVA.
GN      Salmonella typhimurium.
OS      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Salmonella.
OX      NCBI_TaxID=602;
[1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-SL1303;
RX      MEDLINE=95011654; PubMed=7926834;
RA      Hongo E., Morimoto Y., Mita K., Machida I., Hama-Inaba H., Tsuji H.,
RA      Ichimura S., Noda Y.;
RT      "The methyl viologen-resistance-encoding gene smva of Salmonella
RL      typhimurium."
RL      Gene 148:173-174(1994).
-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN, INNER MEMBRANE
(POTENTIAL).
-1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).
CC
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/sib.ch).
CC      or send an email to license@sib-sib.ch).
CC
CC      EMBL: D26057; BAA0505.1; -.
DR      StyGene; SG10384; smva.
KW      Transmembrane; Inner membrane.
FT      TRANSMEM 5 25
FT      TRANSMEM 44 64 POTENTIAL.
FT      TRANSMEM 73 93 POTENTIAL.
FT      TRANSMEM 96 116 POTENTIAL.
FT      TRANSMEM 135 155 POTENTIAL.
FT      TRANSMEM 158 178 POTENTIAL.
FT      TRANSMEM 192 212 POTENTIAL.
FT      TRANSMEM 220 240 POTENTIAL.
FT      TRANSMEM 260 280 POTENTIAL.
FT      TRANSMEM 299 319 POTENTIAL.
FT      TRANSMEM 327 347 POTENTIAL.
FT      TRANSMEM 357 377 POTENTIAL.
FT      TRANSMEM 391 411 POTENTIAL.
FT      TRANSMEM 470 490 POTENTIAL.
SQ      SEQUENCE 496 AA; 52521 MW; 0F1B23C8BD27BCD CMC64;

Query Match 8.4%; Score 92; DB 1; Length 496;
Best Local Similarity 23.0%; Pred. No. 3.4;
Matches 53; Conservative 31; Mismatches 76; Indels 70; Gaps 8;
37 LGATMLDLYVLGLVYALADPTFHYLYIAQKVMFVAVFLVLTIVLEN---LYLF--- 89
||||| : : ||||| ||

```

```

Db      135  IGHVMTLLAIALLLVYSAFTALKGHL-----SLWVSYTLLGALLGLGFIR 242
Qy      90  QHMKLVMPMPVLVIMFINSATVLTITFIACSA-----AVDTSLRGTTPYNQAAAS 144
Db      243  QATSRPMDMLKFLTHRIILSGVAMMTMTLVYGFELLMADELQPVHGISPE-----AG 298
Qy      145  FPACLVMIAGVS-----AFPSYQAMRG----- 167
Db      299  VFMFLVMAASGSGSLAGVLSRLGLVATGSMALSLSPFGIANTDPTSTQOMQAMGLM 358
Qy      168  --VGSNATSONAGVATTCATAPSGAEAAAGSQRHVTLOASMGALRGVS 215
Db      359  ALLGFSASALLA-----STSAIMAAAPAEKAAAGAEIETAYELGAGLGIA 405

RESULT  14
LMPI_EBVC
ID      LMPI_EBVC      STANDARD;      PRT;      404 AA.
AC      P29362;
DT      01-DEC-1992 (Rel. 24, Created)
DT      01-DEC-1992 (Rel. 24, Last sequence update)
DT      01-OCT-1994 (Rel. 30, Last annotation update)
DE      LATENT MEMBRANE PROTEIN 1 (LMP-1) (P63).
GN      BLNF1.
OS      Epstein-Barr virus (strain Cao) (Human herpesvirus 4).
OC      Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OX      Gammaherpesvirinae; Lymphocryptovirus.
RN      NCBI_TaxID=31525;
RP      [1]
RX      MEDLINE=92013956; PubMed=1681026;
RA      Hu L.F., Zabaovsek Y.E.R., Chen F., Cao S.L., Emberg I., Klein G.,
RA      Winberg G.;
RT      "Isolation and sequencing of the Epstein-Barr virus BLNF-1 gene
RL      J. Gen. Virol. 72:2399-2409(1991)."
CC      1- FUNCTION: THE LATENT MEMBRANE PROTEIN HAS TRANSFORMING ACTIVITY.
CC      CC
CC      CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      CC at the European Bioinformatics Institute. There are no restrictions on
CC      CC use by non-profit institutions as long as its content is in no way
CC      CC modified and this statement is not removed. Usage by and for commercial
CC      CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC      CC or send an email to license@isb-sdb.ch).
CC      CC
DR      EMBL: X58140; CAA41148.1;
DR      EMBL: D10059; BAB00948.1;
DR      PIR: J01434; LABECCA.
KW      Transmembrane; Phosphorylation; Transforming protein.
FT      DOMAIN 1 24
FT      TRANSMEM 25 44 POTENTIAL.
FT      TRANSMEM 52 72 POTENTIAL.
FT      TRANSMEM 77 97 POTENTIAL.
FT      TRANSMEM 105 125 POTENTIAL.
FT      TRANSMEM 139 159 POTENTIAL.
FT      TRANSMEM 166 186 POTENTIAL.
FT      DOMAIN 187 404 CYTOPLASMIC.
SO      SEQUENCE 404 AA; 43769 MW; D04536D3B65FF82E CRC64;

Query Match      8.2%; Score 90.5; DB 1; Length 404;
Best Local Similarity 24.2%; Pred. No. 3.7;
Matches 38; Conservative 25; Mismatches 51; Indels 43; Gaps 6;

Qy      21  ASVSALRDPLGCVSRGLGALMLLQVLGLVALLADPTVHLVYPAVGWVFAVFLMTVT 80
Db      8   APPSAPRPPLGPPPLSSSIGLALLLLALLLFLWLYIYWSDM-----TCGALLVLYSFLMLI 63
Qy      81  IYLFNLVLFQ-----LHMKLYMPMPVPLVIMFINSATVLY--ITFIACSAANDLT 129
Db      64  ILLILIFRRDLCPGLGILLMLITV--LLIALNHLGQALVIGIVFI----- 113

```

QY 130 SLRGTREYNQAAAFACLVMIAYGSAFESYQAMR 166
DB 114 -----EGCL--LVEGIMTYFLETLIMR 132

RESULT 15

B3AR_RAT STANDARD; PRT; 400 AA.

AC P26255;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE BETA-3 ADRENERGIC RECEPTOR.
GN ADRB3 OR ADRB3R.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID:10116;
RN [1]
RX MEDLINE:92084710; PubMed:1721063;
RA Muzza P., Revelli J.-P., Kuhne F., Gocayne J.D., McComble W.R.,
RT Venter J.C., Giacobino J.-P., Fraser C.M.;
RT "An adipose tissue-specific beta-adrenergic receptor. Molecular
cloning and down-regulation in obesity."
RT J. Biol. Chem. 266:24053-24058(1991).
RN [2]
RX MEDLINE:92100057; PubMed:1684635;
RA Graneman J.G., Labners K.N., Chaudhry A.;
RT "Molecular cloning and expression of the rat beta 3-adrenergic
receptor."
RT Mol. Pharmacol. 40:895-899(1991).
RN [3]
RX SEQUENCE FROM N.A.
RX MEDLINE:93178631; PubMed:8382630;
RA Bensaïd M., Kaghad M., Rodriguez M., le Fur G., Caput D.;
RT "The rat beta 3-adrenergic receptor gene contains an intron."
RL FBS Lett. 318:223-226(1993).
CC -1- FUNCTION: BETA-ADRENERGIC RECEPTORS MEDIATE THE CATECHOLAMINE-
INDUCED ACTIVATION OF ADENYLATE CYCLASE THROUGH THE ACTION OF G
PROTEINS. BETA-3 IS INVOLVED IN THE REGULATION OF LIPOLYSIS AND
THERMOGENESIS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: WHITE AND BROWN ADIPOSE TISSUES, AND DIGESTIVE
TRACT.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M74716; AAA74470.1; -
DR EMBL: S73473; AAB20702.1; -
DR EMBL: S56481; AAB25520.1; -
DR EMBL: S56152; AAB25521.1; -
DR PIR: A41679; A41679.
DR PIR: S29808; S29808.
DR PIR: A53281; A53281.
DR HSP: P07700; 1DER.
DR GCRDB: GCR_0287; -
DR GCRDB: GCR_0622; -
DR GCRDB: GCR_0732; -
DR InterPro: IPR000276; -
DR InterPro: IPR000681; -
DR InterPro: IPR002233; -
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHHODPSN.

DR PRINTS: PR00563; ADRENGCB3AR.
DR PRINTS: PR01103; ADRENGICR.
DR PROSITE: PS00237; G-PROTEIN_RECIP_F1_1; 1.
DR PROSITE: PS50262; G-PROTEIN_RECIP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; glycoprotein;
FM Multigene family; Phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 36
FT TRANSSEM 37 60
FT DOMAIN 61 69
FT TRANSSEM 70 88
FT DOMAIN 89 108
FT TRANSSEM 109 130
FT DOMAIN 131 152
FT TRANSSEM 153 175
FT DOMAIN 176 200
FT TRANSSEM 201 222
FT DOMAIN 223 289
FT TRANSSEM 290 311
FT DOMAIN 312 323
FT TRANSSEM 324 344
FT DOMAIN 345 400
FT CARBOHYD 8
FT CARBOHYD 26 26
FT DISULFID 107 186
FT LIPID 358 358
FT CONFLICT 234 235
SQ SEQUENCE 400 AA; 43146 MW; D588540C4BACEB13 CRC64;

Query Match 7.9%; Score 86.5; DB 1; Length 400;
Best Local Similarity 26.8%; Pred. No. 7.3;
Matches 55; Conservative 31; Mismatches 70; Indels 49; Gaps 11;

QY 1 MAEPPSK---VSTRSSPAQGAESAVALRDLGFVSRRLGALMLDLVIG-LVWMLIA 56
DB 1 MAPPHKNGSLAFMSDAPTLDPASANTSGLPVMAAALAGALLATVGGNLLVITATA 60
QY 57 DTPHLLPAIGWVWFVAVFLMLVTVIVLENLYFQHLKLVVNP-----WPLYMT 106
DB 61 RTRP-----LQITVTFVTSLATVADLVVGLVMPGATLALGHPPLGATG 106
QY 107 ENI--SATVLYTAFI--ACSAVD-----LTSLR-GTRPNQAAAFACLVMIAYGV 156
DB 107 CELMTSDVCLVTAISIELCLALADRYLATVNPRLIGLVTKRRARA-----VVLWIV 161
QY 157 SAFESY-----QAMRGVSNATSQ 176
DB 162 SATVFAPIKQWNR-VGADAAQAE 185

Search completed: May 2, 2001, 15:11:39
Job time: 79 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2001, 15:09:56 ; Search time 22.67 Seconds
(Without alignments)
1116,758 Million cell updates/sec

Title: US-09-602-597-4
Perfect score: 1101
Sequence: 1 MAEPFSKVSRTSSPAOGAE.....GSOSRYTLQASWGLRGVSP 216

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_15:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mmc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_unclassified:*
- 13: sp Vertebrate:*
- 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	930	84.5	182	4	Q9Y342
2	158	14.4	281	5	001539
3	137	12.4	177	5	Q9VZ71
4	104.5	9.5	152	11	Q9RL07
5	93	8.4	471	2	Q9JVA3
6	93	8.4	610	5	Q44986
7	93	8.4	693	10	Q9LK50
8	92	8.4	471	2	Q9K032
9	91	8.3	208	5	Q21228
10	91	8.3	379	2	Q9K1T9
11	91	8.3	415	2	Q92739
12	90.5	8.2	273	2	Q68480
13	90.5	8.2	461	10	Q9M100
14	90	8.2	401	10	Q40568
15	90	8.2	539	2	Q45577
16	89	8.1	238	14	Q9Y0X5
17	89	8.1	264	2	Q34535
18	89	8.1	417	5	Q9V0K9
19	88.5	8.0	215	2	Q9K8Y7

20	88.5	8.0	244	5	Q20241	Q20241 caenorhabd
21	88.5	8.0	429	2	Q9RV8	Q9RV8 delinococcus
22	88.5	8.0	3005	14	Q56074	Q56074 hepatitis g
23	88	8.0	331	5	Q93869	Q93869 caenorhabd
24	87.5	7.9	229	8	Q9L9P1	Q9L9P1 terebratul
25	87.5	7.9	464	2	Q9RYX1	Q9RYX1 delinococcus
26	87.5	7.9	1192	10	Q9S770	Q9S770 arbidops
27	87	7.9	122	8	Q03762	Q03762 pediatrum
28	87	7.9	160	5	Q9YX8	Q9YX8 drosophila
29	87	7.9	506	5	Q23576	Q23576 caenorhabd
30	86.5	7.9	212	8	Q79613	Q79613 rusaia ros
31	86	7.8	344	2	Q9XDA1	Q9XDA1 thermom
32	86	7.8	406	5	Q44990	Q44990 caenorhabd
33	86	7.8	449	2	Q45082	Q45082 burkholderi
34	85.5	7.8	519	5	Q09592	Q09592 caenorhabd
35	85	7.7	338	2	Q9KX69	Q9KX69 streptomyc
36	85	7.7	358	2	P77568	P77568 pseudomon
37	85	7.7	516	2	Q53508	Q53508 mycobacteri
38	85	7.7	999	9	Q21882	Q21882 bacterioph
39	85	7.7	1083	4	Q9Y666	Q9Y666 homo sapien
40	85	7.7	1147	2	Q9KV54	Q9KV54 vibrio chol
41	84.5	7.7	314	2	Q9UYF8	Q9UYF8 neisseria m
42	84.5	7.7	397	1	Q27735	Q27735 methanobact
43	84.5	7.7	405	2	Q9RJ60	Q9RJ60 streptomyc
44	84.5	7.7	419	2	Q87601	Q87601 streptomyc
45	84.5	7.7	467	6	Q9T96	Q9T96 bos taurus

ALIGNMENTS

RESULT 1
ID Q9Y342 PRELIMINARY; PRT; 182 AA.
AC Q9Y342;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE PLASMO-LIPIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-KIDNEY;
RA Xie B., Durrie R., Sapirstein V.S.;
RT "Molecular cloning of human plasminogen CDNA."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF137386; AAD33060.1;
SQ SEQUENCE 182 AA; 19986 MW; 6F09AA080E2F67E8 CRC64;

Query Match 84.5%; Score 930; DB 4; Length 182;
Best Local Similarity 100.0%; Pred. NO. 2.3e-64;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAEPFSKVSRTSSPAOGAEASVSALRPDLGFVRSRLGALMLQIVLGLVVALADAPY	60
DB	1	MAEPFSKVSRTSSPAOGAEASVSALRPDLGFVRSRLGALMLQIVLGLVVALADAPY	60
QY	61	HLPPYGVWVFAVFLMTVTLFNLVLEQLHMKLYMPWPLVLMIFNISATVLYTAFI	120
DB	61	HLPPYGVWVFAVFLMTVTLFNLVLEQLHMKLYMPWPLVLMIFNISATVLYTAFI	120
QY	61	HLPPYGVWVFAVFLMTVTLFNLVLEQLHMKLYMPWPLVLMIFNISATVLYTAFI	120
DB	61	HLPPYGVWVFAVFLMTVTLFNLVLEQLHMKLYMPWPLVLMIFNISATVLYTAFI	120
QY	121	ACSAVDLTSLGTRPRYNGRAASFFACLYMAYGVSAFFSQARVGCSNATSQMAGG	180
DB	121	ACSAVDLTSLGTRPRYNGRAASFFACLYMAYGVSAFFSQARVGCSNATSQMAGG	180
QY	181	YA 182	
DB	181	YA 182	

```

RESULT 2
ID 001539 PRELIMINARY; PRT; 281 AA.
AC 001539;
DE 01-JUL-1997 (TREMBLrel. 04, Created)
DE 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 14, Last annotation update)
DE 01-MAY-2000 (TREMBLrel. 14, Last annotation update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE CG15211 PROTEIN.
DE CG15211.
OS Drosophila melanogaster (Fruit fly).

Query Match 14.4%; Score 158; DB 5; Length 281;
Best Local Similarity 28.6%; Pred. No. 5, le-05;
Matches 50; Conservative 27; Mismatches 76; Indels 22; Gaps 5;

Y 18 GAENSVSLRPDLGFRVSRIGALMLDLVLWALADTPHYLYPAYGWMFVAVFLW 77
D 104 GPAVRIEPRDLCEYIRLIGIMKIVICVLCCLLFIFVPMGPAY-YGVGWAFTVSSVGI 162
Y 78 LVTIVLFLYLFOLMKIKYMPPLVY-----MIRINATYLYITAFACSAAY 126
D 163 FVTISLTLFLYFRVVDLPSTINWVCLSTREITIIIDEMVYCFAMTVFFFLA--ACVLAV 220
Y 127 DLTSLAGTRPNORAAAFACLVIAVGVSAFYSQAMRG-----GSMNATSQ 176
D 221 ASSDGRGFAM---AIAFAFGACAGFPCYIKAFLSMKNNERATGGSNPNVIO 272

RESULT 3
ID 09VZ71 PRELIMINARY; PRT; 177 AA.
AC 09VZ71;
DE 01-MAY-2000 (TREMBLrel. 13, Created)
DE 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE CG15211 PROTEIN.
DE CG15211.
OS Drosophila melanogaster (Fruit fly).

```

```

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;

RE SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
AD Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
AN Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan R.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,
RA Adair J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,
RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodes A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Leal Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K.A., Nusskern D.R., Paclab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissensbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003484; AAF47955.1; -.
DR FLYBASE: FBgn0030234; CG15211.
SQ SEQUENCE 177 AA; 10910 MW; ED40DBA6B8BD1A1 CRC64;

Query Match 12.4%; Score 137; DB 5; Length 177;
Best Local Similarity 27.1%; Pred. No. 0.0014;
Matches 45; Conservative 29; Mismatches 78; Indels 14; Gaps 4;

Y 9 STRTSSPAQGAENSV-----SALRPDLGFRVSRIGALMLDLVLWALADTPHY 61
D 13 TTTTNGPPGANGPNVSGGFGFWSAIRINIDFFRIPIGIKIVEVLAIGICMAAP--- 69
Y 62 LYPAYGVMMFVAVFLVLTIVLFLYLFOLMKIKYMPPLVY-----MIRINATY 120
D 70 LASATSEFMFVYIVSIFILINLLIAYLGLREALNVVMNMFSELTITVAVTLTYLFTGFI 129
Y 121 ACSNAVDLTSLAGTRPNORAAAFACLVIAVGVSAFYSQAMR 166
D 130 VQLRMSDADATGKS---GSNTAGVDFGLFNLAIYAAGTYFLFLHR 172

RESULT 4
ID 09RI07 PRELIMINARY; PRT; 152 AA.
AC 09RI07;
DE 01-MAY-2000 (TREMBLrel. 13, Created)

```

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE PROTEOLIPID PROTEIN 2.
 GN PLP2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hayashi A., Hattori A., Okaze H., Kozuma S., Seki N., Satto T.;
 RT "Mouse proteolipid protein 2 (PLP2)."
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB031292; BAA83500.1; -
 SO SEQUENCE 152 AA; 16607 MW; C651A5E43B2909D6 CRC64;

Query Match 9.5%; Score 104.5; DB 11; Length 152;
 Best Local Similarity 22.9%; Pred. No. 0.38;
 Matches 30; Conservative 26; Mismatches 60; Indels 15; Gaps 3;

OY 32 EVRSRLGALMLQLVGLLWALLADTPYHLXPAYGMVAVFLVTVILENLYLFEOL 91
 DB 19 FSRTKGIILFAEIIICLVILICFSASTSAYSS-----LSVTEMICAAVLVFTCDL 72
 OY 92 HMKLVVPPVPLVIMINISATVYITAFACSAVDLSRGTRPTNORAAAFPAQVW 151
 DB 73 HSKISFIMPWDFRSLATILYLTSTI-----VVLVEGRG-----SRVVAIGILGLL 123
 OY 152 IAVGSAFFSY 162
 DB 124 LFGYDAITTF 134

RESULT 5
 O93V43 PRELIMINARY; PRT; 471 AA.
 AC O93V43:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE PUTATIVE TRANSMEMBRANE TRANSPORT PROTEIN.
 GN NMA1003.
 OS Neisseria meningitidis (serogroup A).
 OC Bacteria; Proteobacteria; Beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=63699;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=22491 / SEROTYPE 4A;
 RX MEDLINE=20222556; PubMed=10761919;
 RA Parkhill J., Achtman W., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holtroyd S.,
 RA Jags R., Leather S., Moule S., Mungall K., Quail M.A.,
 RA Raine A., Rutherford K.M., Simmonds M., Skelton J.,
 RA Whitehead S., Spratt B.G., Barrett B.C.;
 RT "Complete DNA sequence of a serogroup A strain of Neisseria
 meningitidis 22491."
 RL Nature 404:502-506(2000).
 DR EMBL: AL162754; CAB84272.1; -
 KM Transmembrane.
 SO SEQUENCE 471 AA; 50623 MW; 1E0A3A08B8F6F6CB CRC64;

Query Match 8.4%; Score 93; DB 2; Length 471;
 Best Local Similarity 19.4%; Pred. No. 7.7;
 Matches 42; Conservative 37; Mismatches 74; Indels 64; Gaps 9;

OY 39 ALMLDLVGLL-----VVAL-----IADTPYHLTPA----- 65
 DB 165 AAMMPLANGM.SHLDEREKRYFVLLGIACASIGIGLTVGSPNDIAAKALNLD 224
 OY 66 YGVV-MEVAVFLVTVIVFNLYL-----QLHMKLVVWPPL-----VLMFINISATV 113

DB 225 VGMKRLGLPMMLLILPLMLSLVILAKPNLNEVEIKASIPWTLHRYALALLIFATAA 284
 OY 114 LYTFACSAVDLTLSTGRTPNORAAAFACIAMIATVSAFFSYOARGVGSN-- 171
 DB 285 WIPSSKIKTAFGIS-----NPDVIALSAVAVVFVA-----QKEVAARN 328
 OY 172 -AATSGMAGYATCATAPSGAANAAGSOSRYTLAS 207
 DB 329 WGYLMFGGISLSTLKTSGSEALGQVAAATFSGA 365

RESULT 6
 O44986 PRELIMINARY; PRT; 610 AA.
 AC O44986:
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE K02F2.6 PROTEIN.
 GN K02F2.6.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Faveille A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkneen R.,
 RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Spoat J., Woilman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.";
 RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Magg L., Goela D.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL: AF043699; AAB97568.1; -
 DR INTERPRO: IPR000276; -
 DR PFM; PFO0001; 7tm_1.1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G_PROTEIN_RECEPTOR; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein.
 SO SEQUENCE 610 AA; 67762 MW; 57DB24D638566C21 CRC64;

Query Match 8.4%; Score 93; DB 5; Length 610;
 Best Local Similarity 22.8%; Pred. No. 9.6;
 Matches 42; Conservative 39; Mismatches 67; Indels 36; Gaps 10;

OY 33 VSRRLGALMLQLVGLLWALLADTPYHLXPAYGMVAVFLVTVILENLYLFEOL 84
 DB 89 LRSATLILSLALADLGLTVLPISIANEVLDQWIGETWCTIWLIDIMCTASTY 148
 OY 85 NYLFLQHLKLVWP---WPLVLMINISATV---LYTFACSAVDL-----TSL 131

DB 149 NVAISIDRYIAIKPLNPLMTKFRACVTAIYWGISFLICSPSFLASSIKOKETPC 208
 QY 132 KGRIPNO-----RAAASFACVMTAIGVSAFSSQAMRGVSNATSMAGCIATT 184
 DB 209 KCT-PANAGRYVVSASSSFYPMIYF---VYFRIT---VAARATKSIYSIGMSV 260
 QY 185 CATA 188
 DB 261 TAAA 264

RESULT 7

Q9LK50 PRELIMINARY: PRT: 693 AA.

AC 09LK50: 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DE ABC TRANSPORTER-LIKE PROTEIN.
 OS Arabidopsis thaliana (mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RA Nakamura Y.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
 RT Sequence features of the regions of 4,251,695 bp covered by ninety p1,
 RT TAC and BAC clones."
 RL DNA Res. 7:217-221(2000).
 DR EMBL: AP000375; BAB01414.1;
 SO SEQUENCE 693 AA; 78940 MW; 48F8D87413209A2 CRC64;

Query Match 8.4%; Score 93; DB 10; Length 693;
 Best Local Similarity 19.6%; Pred. No. 11;
 Matches 32; Conservative 35; Mismatches 72; Indels 24; Gaps 4;

QY 32 FVRSRLGALMLQVLYGLVVALADPYHLDPAYGWVFAVFLMTYIYFNLFLPOL 91
 DB 440 FQKRLVQSLGVAAYVGLMMKSKDTDEHLRDQGLMFYICIF-WTSSSLFGAYV-- 496
 QY 92 HMKLWVWPLVLMFNISATVLYTATLACSAVDLT-----LGRTRPVNOR 140
 DB 497 -----PFKKLYLVKERRAKETRLSYVYVCSLCDMAVHLYPTFEMIYFMAEFNRN 549
 QY 141 AAASFACVMTAIGVSAFSSQAMRGVSNATSMAGCIATT 183
 DB 550 IPCFLFVLTITLITAIT---SOGAGEFTGASVLSIKRGMATAS 589

RESULT 8

Q9K032 PRELIMINARY: PRT: 471 AA.

AC 09K032: 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DE 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE TRANSPORTER, NADCO FAMILY.
 GN NMB0792.
 OS Neisseria meningitidis (serogroup B).
 CC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MC58 / SEROGROUP B;

RX MEDLINE=20175755; PubMed=10710307;

RA Tetteijn H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 RA Hatt D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Clecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
 RA Gill J., Scarlato V., Maignan V., Pizsa M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 MC58".
 RL Science 287:1809-1815(2000).
 DR EMBL: AE002433; AAF41205.1;
 DR TIGR; NMB0792;
 SO SEQUENCE 471 AA; 50647 MW; 2DBC53ECC5ICE04 CRC64;

Query Match 8.4%; Score 92; DB 2; Length 471;
 Best Local Similarity 19.4%; Pred. No. 9.2;
 Matches 42; Conservative 37; Mismatches 74; Indels 64; Gaps 9;

QY 39 ALMLQVLYGLL-----VVAL-----IADPYHLVPA----- 65
 DB 165 AAMPLPLMGLSHDQKEHKTYVFLGLGAYCASIGSGTGLVSPPLIAKALNLD 224
 QY 66 YGVV-MEVAVFLMLVTVLFNLFLF-----QLMKLWVWPL-----VLMFNISATV 113
 DB 225 VGMKMLGIPMMLLIPMLLSLYVLKPNLNERVEIKESISMTLHTRYALILFLATAA 284
 QY 114 LYTRAFICSAVAVDLSLGRTRPVNQRRAASFACVMTAIGVSAFSSQAMRGVSN-- 171
 DB 285 WIFSSKITAFGIS-----NPDVIALSAVAVVFGVA-----QMKVARNRD 328
 QY 172 -AATSQMAGYATTCATAPSGAAGAAAGSRTVLTAS 207
 DB 329 GGVMLFPGGIGISLTLTKTSGNSALGQVATFSGA 365

RESULT 9

Q21228 PRELIMINARY: PRT: 208 AA.

AC 021228: 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DE 01-JAN-1999 (TREMblrel. 09, Last annotation update)
 DE K04G2.9 PROTEIN.
 GN K04G2.9
 OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Gardner A.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Parsons J., Percy C., Riffen L., Roodra A., Mortimore B., O'Callaghan M.,
 RA Saldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weissstock L., Wilkinson-Sproat J., Wohlman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans".
 RL Nature 368:32-38(1994).
 DR EMBL: Z75712; CAB00046.1;
 SO SEQUENCE 208 AA; 23669 MW; 4EB921774853A61F CRC64;

Query Match 8.3%; Score 91; DB 5; Length 208;
 Best Local Similarity 24.0%; Pred. No. 5.4;
 Matches 45; Conservative 30; Mismatches 76; Indels 32; Gaps 8;

QY 8 VSTRSSPAAGASASAL---RPDLGFVSRGALMLDLVGLVMAIADTPHYLP 64
 DB 17 VTTTTRKTKYKHSVWSCGHPDKNYCLGPIGLRIAIFLCVIVCLITSV-FGPGP 75
 QY 65 AYGWVAF-----VAVFLVTVIVLNFYLPOLHMK-LYVWPPLVIMFISATVLYI 116
 DB 76 FRG-IIFGQTILTVASVAMLLFFILIAFFTLHSHDFPCRRDILFNYICAVLFI 134
 QY 117 TAFIA-----CSAAYDLTS---LQSTPRYNORAASEFACVMAVGSAAF 160
 DB 135 VLSIVAYYSTGSMNKN---DIGSDGIHNCRIITWEAFASFILFIIGIFALTAFL 191
 QY 161 SYQ 163
 DB 192 THK 194

RESULT 10

Q9K1T9 PRELIMINARY; PRT; 379 AA.
 AC Q9K1T9;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE CELL SHAPE-DETERMINING PROTEIN MRDB.
 GN CP1002.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=83558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AR39;
 RC MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Umayam L.A., Utterberg T.,
 RA Berry K., Bass S., Linher K., Weidman J., Khouri H., Craven B.,
 RA Bowman C., Dodson R., Gwin M., Nelson J., Deboy R., Kolonay J.,
 RA McCarty G., Salzberg S.L., Eisen J., Fraser C.M.,
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
 RT pneumoniae AR39."
 RL Nucleic Acids Res. 28:1397-1406(2000).
 DR EMBL; AE002257; AAF38780.1;
 DR TIGR; CP1002;
 SO SEQUENCE 379 AA; 42445 MW; 424D84D2317F14F CRC64;

Query Match 8.3%; Score 91; DB 2; Length 379;
 Best Local Similarity 24.0%; Pred. No. 9;
 Matches 35; Conservative 24; Mismatches 49; Indels 38; Gaps 5;

QY 23 VSALRPDLGVRRLGAL-----MLQVLGLVMAIADTPHYLPAYGVAFVAF 76
 DB 29 ISSMDPRAMLYTSKGLITKSTIMOLRHPALGWNVPFICVYHLEKRAVLY--FFM 86
 QY 77 MLVTIVLFNLYLFOLHMKLYVFW-----PLVIMFISATVLYTAFIACSAA 125
 DB 87 ICALVGLGFVPSVQVNHVRYRIPFIHMSVQSEYKLVIVY-----MLSYILES 137
 QY 126 VDLTSLRGTRPYNORAASEFACVMA 151
 DB 138 ADITS-----KTTAFLACTLV 153

RESULT 11

Q92739 PRELIMINARY; PRT; 415 AA.
 ID Q92739;
 AC Q92739;

DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE ROD SHAPE PROTEIN.
 GN RODA.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=83558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CWL029;
 RC MEDLINE=99206606; PubMed=10192388;
 RA Kalman S., Mitchell W., Marathe R., Lamell C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.,
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."
 RL Nat. Genet. 21:385-389(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-J138;
 RC MEDLINE=20330349; PubMed=10871362;
 RA Shirai M., Hirakawa H., Kimoto M., Taduchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.,
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RT from Japan and CWL029 from USA."
 RL Nucleic Acids Res. 28:2311-2314(2000).
 DR EMBL; AE01667; AAD19005.1;
 DR EMBL; AP002548; BAA99075.1;
 DR INTERPRO; IPR000255;
 DR INTERPRO; IPR001182;
 DR PIRAM; PF01098; FTSW RODA.SPOVE; 1.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
 DR PROSITE; PS00428; FTSW RODA.SPOVE; UNKNOWN_1.
 SO SEQUENCE 415 AA; 46662 MW; 23222066F5A6B5C CRC64;

Query Match 8.3%; Score 91; DB 2; Length 415;
 Best Local Similarity 24.0%; Pred. No. 9.8;
 Matches 35; Conservative 24; Mismatches 49; Indels 38; Gaps 5;

QY 23 VSALRPDLGVRRLGAL-----MLQVLGLVMAIADTPHYLPAYGVAFVAF 76
 DB 65 ISSMDPRAMLYTSKGLITKSTIMOLRHPALGWNVPFICVYHLEKRAVLY--FFM 122
 QY 77 MLVTIVLFNLYLFOLHMKLYVFW-----PLVIMFISATVLYTAFIACSAA 125
 DB 123 ICALVGLGFVPSVQVNHVRYRIPFIHMSVQSEYKLVIVY-----MLSYILES 173
 QY 126 VDLTSLRGTRPYNORAASEFACVMA 151
 DB 174 ADITS-----KTTAFLACTLV 189

RESULT 12

O68480 PRELIMINARY; PRT; 273 AA.
 ID O68480;
 AC O68480;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE ACCESSORY PROTEIN NOST PRECURSOR.
 GN NOST.
 OS Achromobacter cycloclastes.
 OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 OC Achromobacter.
 OX NCBI_TaxID=223;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 21921, IAM 1013;
 RA McGuire M.A., Nelson L.K., Bollinger J.A., Chan Y.-K., Dooley D.M.,
 RL J. Inorg. Biochem. 0:0-0(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-IAM 1013;

RX MEDLINE-99156231; PubMed-10048486;
 RA Inatomi K.;
 RT Analysis of the nitrous oxide reduction genes, nosZDYL, of
 RL Achromobacter cycloclastes.;
 DR DNA Ref. 5:365-371(1998);
 DR EMBL; AF047429; RAD09160.1.;
 KW EMBL; Y15161; CAA75428.1.;
 FT Signal.
 SQ SEQUENCE 273 AA; 27648 MW; 1CDAB572876E28ID CRC64;

Query Match 8.2%; Score 90.5; DB 2; Length 273;
 Best Local Similarity 30.8%; Pred. No. 7.4; Mismatches 56; Indels 33; Gaps 7;

OY 21 ASVSALRPLDGFVRSRIGALMLQVLGLVWALADPYHLV-----AYGVWFEVA 73
 DB 123 ALAAVWFPDQ-ATAGLGALML-IMSSVLGATFLGTCYALSALARRPSGAG----LA 176
 OY 74 VELWLVTVLENYLFEQL-----HMKLYVFWPLV-----LMFNISATVLYTAFI 120
 DB 177 VALWLVAVVLYDLALALVTDGGAFTTHALPALLANPADAFRFFNLAAQ----- 229
 OY 121 ACSAAVDLTSLRTPRYNORAAAFACVIMAYGVSAF 159
 DB 230 AVSAGGLGGAAGTIFLWOSASLAMPALATLAAAF 268

RESULT 13

OSM100 PRELIMINARY; PRT; 461 AA.

AC 09M100; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE PUTATIVE SUGAR TRANSPORTER.
 GN AT4604750
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lamer B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,
 RL Mayer R.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL161501; CAB80840.1.;
 KW Sugar transport.
 SQ SEQUENCE 461 AA; 49470 MW; 652408EDD93DA3CA CRC64;

Query Match 8.2%; Score 90.5; DB 10; Length 461;

Best Local Similarity 22.9%; Pred. No. 12;
 Matches 46; Conservative 28; Mismatches 62; Indels 65; Gaps 10;

OY 1 MAE-----FBSKSTRSSPAGQ-----AEASVSALRP-----DLGFVSRGAL 40
 DB 1 MAEESILPHEBTDVASPKRSSLSLSISNASTRPFVAFVSGGALSPFCISFFGSI 60
 OY 41 MLQVLGLVWALADPYHLVPGVWFAVEL-RLVTVLENYLFLQHKML----- 95
 DB 61 LTVGLLGLGKRLADLVGRYTT--WITNVLVIGWALAFADVALLDLGRLLGIGIS 118
 OY 96 -----YMPWPLVLMIFNISATVLYTAFIACSAVDLTSLRTPRYNORAAAFACV 150
 DB 119 VGSSVYLP-----YISEL-----APRLRGAASSLQGL- 148

OY 151 MIAGVSAPFEESY---QAMGV 168
 DB 149 FVGGLSAFYALGTAVAMRSL 169

RESULT 14

ID 040568 PRELIMINARY; PRT; 401 AA.

AC 040568; 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE PHOSPHATE TRANSLOCATOR PRECURSOR.
 OS Nicotiana tabacum (common tobacco).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 OC Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=LEAF;
 RC MEDLINE-94166771; PubMed-8121415;
 RX Knight J.S., Gray J.C.;
 RT "Expression of genes encoding the tobacco chloroplast phosphate
 RL translocator is not light-regulated and is repressed by sucrose."
 RL Mol. Gen. Genet. 242:386-394(1994).
 DR EMBL; X75088; CAA52979.1.;
 DR MENDEL; 9334; Nicta;1243;9334.
 KW Signal.
 FT SIGNAL 1 73
 FT CHAIN 74 401
 FT SEQUENCE 401 AA; 43705 MW; 36CC65B59A945035 CRC64;

Query Match 8.2%; Score 90; DB 10; Length 401;

Best Local Similarity 22.2%; Pred. No. 11;
 Matches 48; Conservative 31; Mismatches 73; Indels 64; Gaps 10;

OY 10 TRTSSPAGASVSALRPDLGFVRSRIGALMLQVLGLVWALADPYHLVPGVW 69
 DB 72 TAASSPAGSDSADDA-----KVGFFNKATLNGFF----- 102
 OY 70 MEVAVELMLVTVLENYLFEQLHKLY-MVWPLVLMFNISATVLYI-----TAFIACA 124
 DB 103 -----FEMWYELNVLENI-----LNKRYNPFYPPVSVIHLAVGYCLISWTGLPRA 154
 OY 125 AVDLTSLRTPRY-----NORAAAFACVIMAYGVSA-----FFSQAMRGVGSNAT 174
 DB 155 PIDSTOLKLLPVAFCALGHVTSVNSFAVAVSTHTIKALEPF-----NASA 204
 OY 175 SOMAGYATTCATAPSGAAG--SOSRYT-LOASW 208
 DB 205 SQFLLGQOIPALMLSLAPVVLGVSMASLTLSFWM 240

RESULT 15

Q45577 PRELIMINARY; PRT; 539 AA.

AC 045577; 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 06, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE YBEC PROTEIN (ORF3).
 GN YBEC.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-168;
 RX MEDLINE-98044033; PubMed-9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

RA Azevedo V., Bertero L.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brynneil S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.U., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Enrich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabre C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Gallizi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grand G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C., Hentut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullio M.F., Itaya M., Jones L.,
RA Jolis B., Karamata D., Kasahara Y., Kluerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kunano M.,
RA Kunita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
RA Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Priesean E., Puig P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schwoert R., Scofield F.,
RA Setiguchi J., Sekowska A., Serró S.J., Serró P., Shin B.S., Soldo B.,
RA Sotokhin A., Tacconi E., Takai T., Takahashi H., Takemaru K.,
RA Takeuchi M., Yamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Totsato V., Uchiyama S., Vandenbol M., Vanlier F., Vassarelli A.,
RA Viari A., Mambrut R., Wedler H., Wedler H., Weltzenegger T.,
RA Vinters P., Wipit A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zunshteln E., Yoshikawa H., Danchin A.,
RA "The complete genome sequence of the gram-positive bacterium *Bacillus*
RT *subtilis*." ;
RT *subtilis*." ;
RL Nature 390:249-256(1997) .
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN-168;
RC STRAIN-168;
RC Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 526-539 FROM N.A.
RC STRAIN-BR95;
RC MEDLINE-94282317; PubMed-8012593;
RX Nilsson R.P., Beller L., Rutberg B.;
RT "The glpT and glpQ genes of the glycerol regulon in *Bacillus*
RT *subtilis*." ;
RL Microbiology 140:723-730(1994) .
RN [4]
RN SEQUENCE FROM N.A.
RP STRAIN-168;
RC Haga K., Liu H., Yasunoto K., Takahashi H., Yoshikawa H.;
RA "Sequence analysis of the 70kb region between 17 and 23 degree of the
RT *Bacillus subtilis* chromosome." ;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z99105; CAB12006.1; -
DR EMBL; Z26522; CAB81293.1; -
DR EMBL; AB006424; BAA33109.1; -
DR INTERPRO: IPR002027; -
DR INTERPRO: IPR002293; -
DR Pfam; PF00324; aa_permeases; 1.
SO SEQUENCE 539 AA; 59197 MW; 2D022FB15CFA0610 CRC64;

Query Match	8.2%	Score 90;	DB 2;	Length 539;
Best Local Similarity	28.3%	Pred. No. 15;		
Matches	39;	Conservative	19;	Mismatches 36;
				Indels 44;
				Gaps 9

```

0Y  62  LYPANGMWFAVAVLEMTLYTLFNLTLFEOLHMKIKYMPWPLMLTFMISANTVYITFA 122
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  335  LHPITG-VPROAMFENLVSTF-LFTLR-----GQGLAEIISVAFLISIT 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
0Y  122  CSAAYDLTSLRGT-----RP-----YNQRAASF-FACLVN-----IAYGV 156
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
,Db  381  --GPIVTTLRRTKDKLYRPLRLKGLNVINAPLGIFIASVLVYMARWPLTGOVLEIILGL 438
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
0Y  157  SAFESYCA--WRGVGSN 171
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Db 439 PIFYYQAKAKWKGFSRN 456

Search completed: May 2, 2001, 15:11:27
Job time: 91 sec
